VOLUME 106 PART B NUMBER 29 SEPTEMBER 1959

The Proceedings

OF

THE INSTITUTION OF ELECTRICAL ENGINEERS

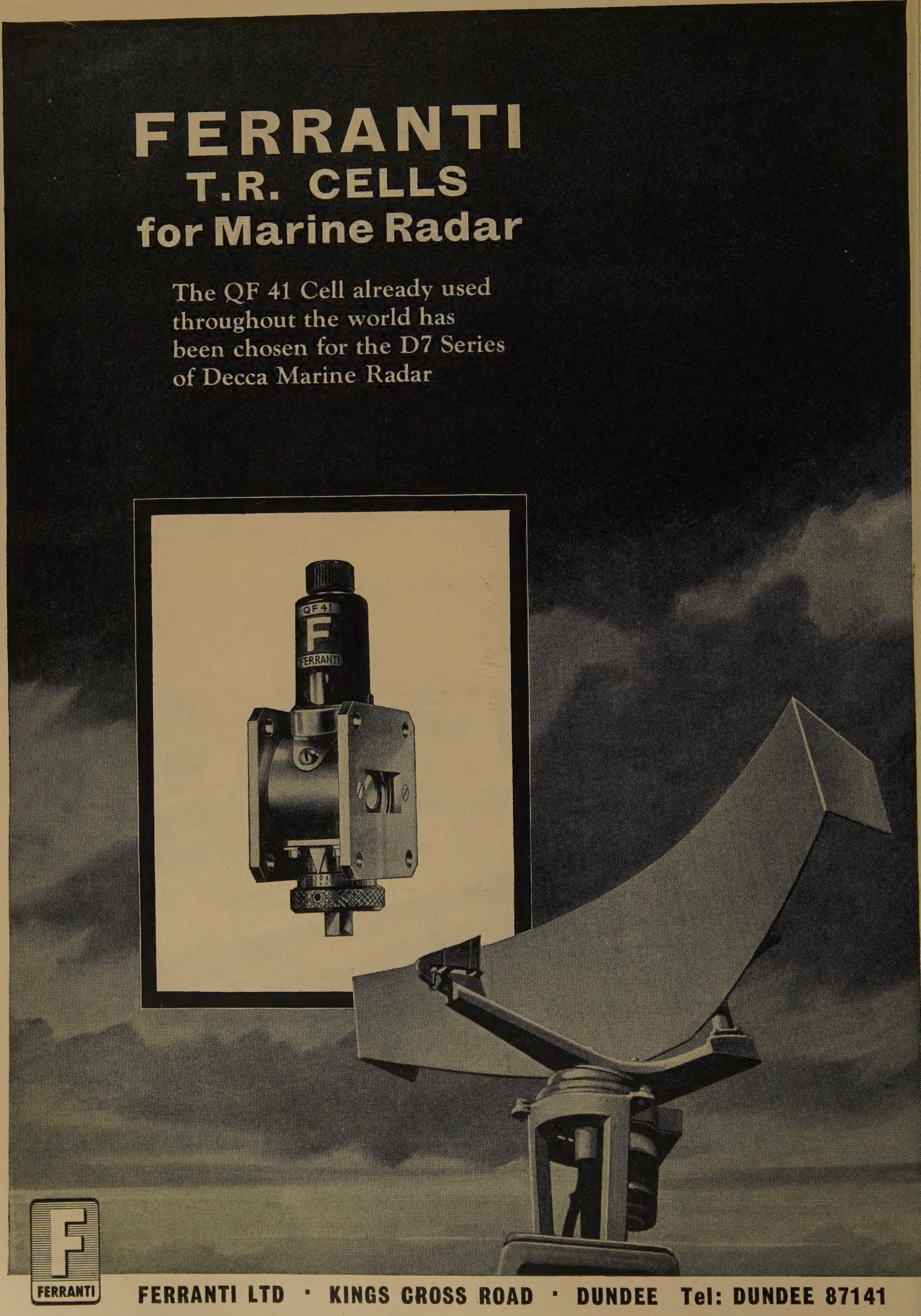
FOUNDED 1871; INCORPORATED BY ROYAL CHARTER 1921

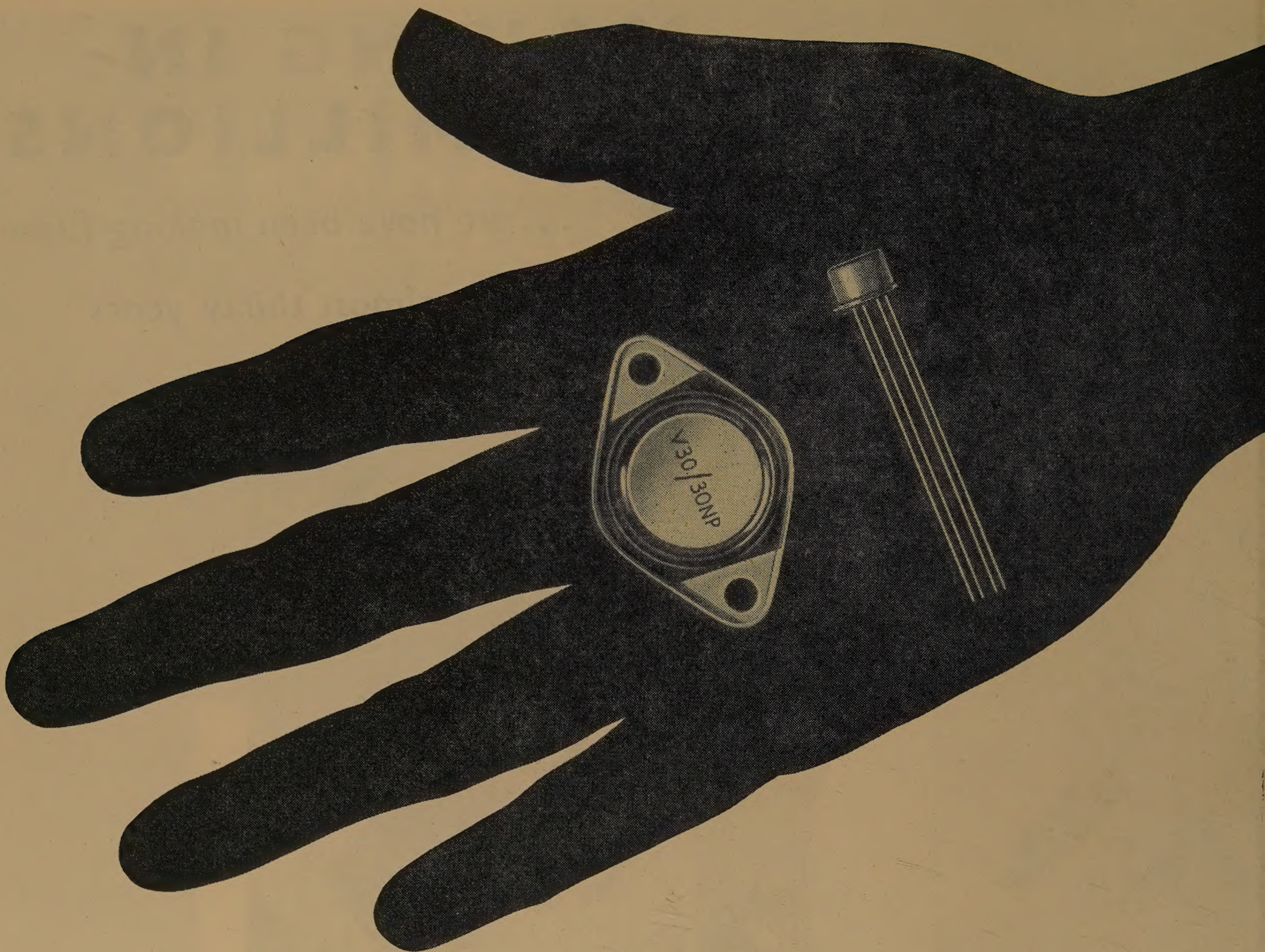
PART B

ELECTRONIC AND COMMUNICATION ENGINEERING
(INCLUDING RADIO ENGINEERING)

SAVOY PLACE · LONDON W.C.2

Price Ten Shillings and Sixpence


FERRANTI
T.R. CELLS
for Marine Radar
The QF 41 Cell already used
throughout the world has
been chosen for the D7 Series
of Decca Marine Radar
QF 41
F
FERRANTI
F
FERRANTI
FERRANTI LTD · KINGS CROSS ROAD · DUNDEE Tel: DUNDEE 87141



Which of these could solve your problem?

Newmarket Transistors Limited have added the following transistors to their range:—

The new Noodle Transistor

(V15/15NP V15/30NP V30/15NP V30/30NP)

1. Amplifies at 10 amps.
2. American diamond construction interchangeability.
3. No fixing screw problems.
4. Cold welded case for increased reliability.
5. Higher frequency operation.

V.H.F. Drift Transistor

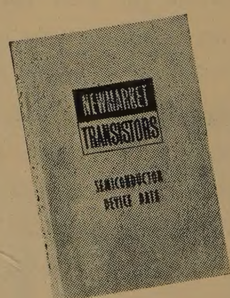
(V15/20 OR)

1. Oscillates up to 100 M/Cs.
2. Amplifies up to 20 M/Cs
3. Jedec 30 case.
4. Higher operating voltage.

Fast Switching Transistor

(V10/1S V10/2S)

1. Phenomenal peak current performance—will switch 1 ampere at 1 amp. peak.
2. Fifty Milli-microseconds risetime.
3. 20V collector rating.
4. British standard cylindrical construction.
5. Cold welded case for increased reliability.



Send for data sheets of these transistors. Ask also for copy of our booklet "Semiconductor Device Data" if you do not already have one.

NEWMARKET TRANSISTORS

Newmarket Transistors Limited.

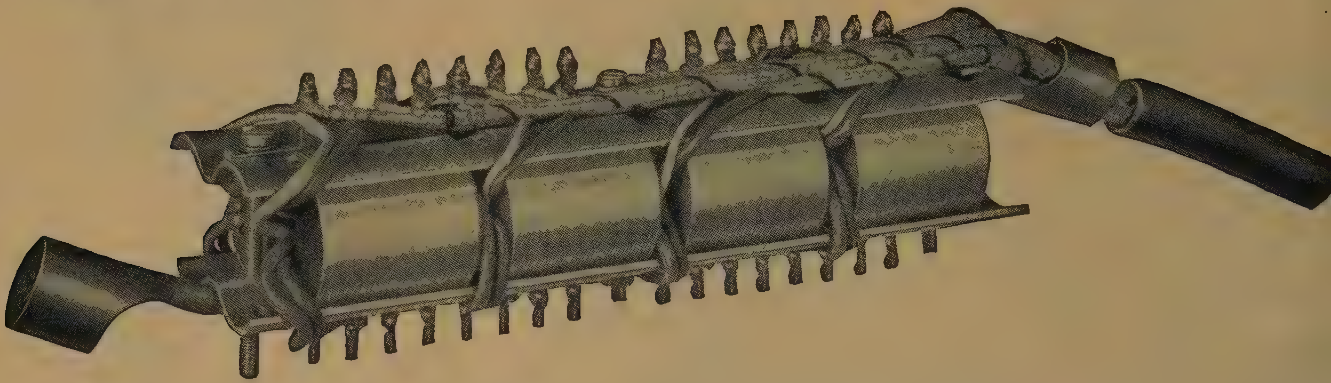
EXNING ROAD, NEWMARKET, SUFFOLK.

Telephone: Newmarket 3381/4 Cables: Semicon Newmarket.


HAMILTON
NORTH ISLAND
ŪRENUI
NEW PLYMOUTH
PALMERSTON NORTH
NEW ZEALAND
NORTH ISLAND
SOUTH ISLAND
4000 MC/S 600-CIRCUIT CAPACITY 2 WAY TERMINAL AND REPEATER
0 10 20 40 60 80 MILES
0 10 20 40 60 80 100 120 KILOMETRES


VHF
CANADA
HAITI
VENEZUELA
G.E.C.

Improved splice loading with the L219



The new economical loading coil

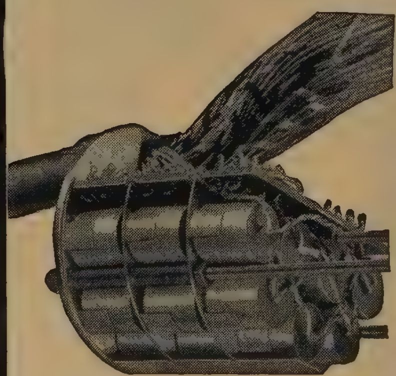
Arising from the increasing demand for a smaller coil which can be employed in splice loading, the L219 has been developed. In the design Mullard Equipment Limited were assisted by their own production experience and information given by overseas users. The result is a simple, low cost component (to grade 3 spec.) suitable for small or large splice loading units.

Smaller construction

By using a new grade of Ferroxcube pot core the overall volume of the coil is considerably reduced. The coil is resin sealed in a small cylindrical aluminium canister ensuring complete protection from climatic effects. The windings of the coils are brought out on flying leads.

LIFE SIZE COIL
L219

Permits smaller splices



Key factors in this development are the clamping arrangements which, with the new coil, permit much smaller splice housing. On small cables, coils are mounted lengthways in pairs with great compactness. For larger cables, coils are mounted radially, each mounting plate accommodating up to seven coils. Clamping plates, coils, etc. can be supplied as kits.

Please write for full details of these new loading coils

MULLARD EQUIPMENT LIMITED

A Company of the Mullard Group

Mullard House · Torrington Place · London W.C.1 · Telephone: Langham 663

ME61


RADIO
CHANNELLING
EQUIPMENT R24B
• For up to 24 carrier channels
• Fully transistorised and compact
• Advanced components, construction and design
• C.C.I.T.T. performance
• Inbuilt signalling equipment, dialling or ring down
This equipment combines up to 24 speech circuits, in two groups of twelve, into a composite signal for feeding into suitable v.h.f. and u.h.f. transmission systems. A complete twenty-four channel terminal occupies one bayside. For further details request leaflet. TEL. 3301
ATE
TRANSMISSION EQUIPMENT TYPE CM.
AUTOMATIC TELEPHONE & ELECTRIC CO. LTD
STROWGER HOUSE, ARUNDEL STREET, LONDON, W.C.2

AT 8831

The best 75 volt stabilisers in the world

BRITISH SERVICES PREFERRED TYPE

M8225/CV4080

The high performance of the Mullard stabiliser 75C1 has led to the recent adoption of its Special Quality equivalent M8225/CV4080 by the British Services as their Preferred 75-volt stabiliser. The M8225/CV4080 is tested for specialised applications in which conditions of extreme shock and vibration are encountered.

- **Wide Current Range . . .**
  2 to 60 milliamps
- **Small Regulation Voltage . . .**
  Less than 9 volts
- **High Stability . . .**
  Typical variation in burning voltage less than ± 2% in any 10,000 hours of operation.

GENERAL PURPOSE TYPE

75C1

The 75C1 is the best 75 volt stabiliser available in the world for general purpose use in industry and communications. It has the same electrical characteristics as the M8225/CV4080 and like this British Services Preferred valve provides an exceptional *combination* of long life, stability and good regulation.

Full data is readily available from the address below.

MULLARD LIMITED
MULLARD HOUSE
TORRINGTON PLACE
LONDON · W.C.I
TEL: LANGHAM 6633

Mullard
GOVERNMENT AND
INDUSTRIAL VALVE DIVISION

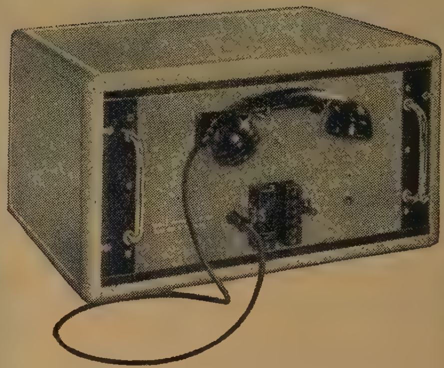
MVT355

TRANSISTORISED

1 + 1 CARRIER SYSTEM E.T.L.14

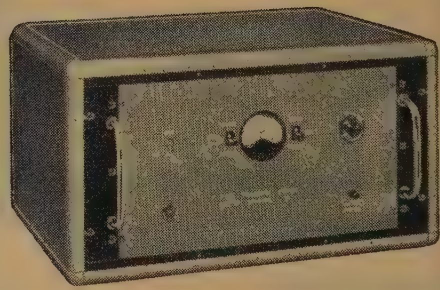
This single channel carrier system is used over an existing two-wire line to provide an extra speech circuit between 3.3 kc/s and 8.7 kc/s. The equipment will operate satisfactorily without repeaters over some 400 miles of open wire line. Signalling over the carrier circuit is at 500 c/s equivalent, standard 17 c/s ringing is used between the carrier terminal and the switchboard line.

Each terminal can be supplied in a metal cabinet or in rack mounting form. Power consumption is less than 0.25 watts and the system can be energised from mains, trickle charged accumulator or dry batteries.



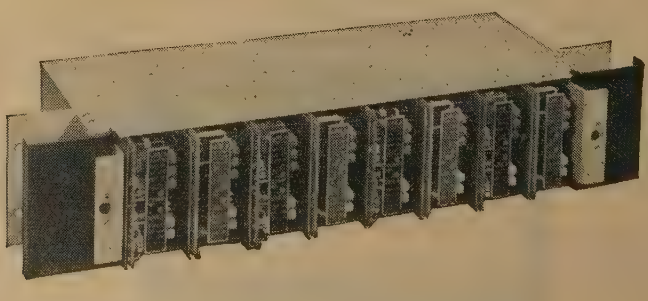
F.M.T.17

The F.M.T.17 is intended for use over existing audio circuits. It enables frequency modulated duplex telegraph signals to be transmitted simultaneously with normal speech, without mutual interference and without any increase in bandwidth. Telegraph speeds are sufficiently high for good teleprinter operation and facilities are provided for single and double current working. The equipment is of unit construction and can be supplied for rack mounting or in a metal cabinet. Power requirements are very modest and operation can be from 12 or 24 volts D.C. or direct from mains.



COMPANDORS

The compandor offers an economical means of improving performance on circuits which suffer from induced noise or crosstalk. It consists of two units, the compressor which operates in the transmit path, compressing the dynamic range in the ratio 2 : 1, low level signals being raised with respect to high levels and the expander which is fitted in the receive path to restore the relative levels. Overall insertion loss is O db. Power is obtained from 12 volt batteries or suitable mains units; only 0.12 watts is required per compandor. The equipment is suitable for panel mounting and a standard $3\frac{1}{2}''$ deep panel will accommodate eight units, i.e. four compressors and four expanders.



transmission equipment

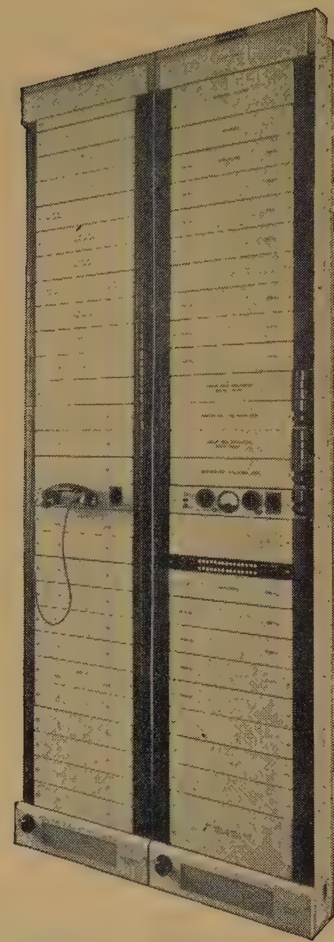
Ericsson Transmission Equipment is being developed continuously to meet the requirements of all administrations. Recent developments include the following fully transisterised items—

Multi-channel Carrier Termination equipment for use over radio and cable circuits.

A frequency modulated telegraph system enabling telegraph signals to be transmitted simultaneously with speech over existing telephone networks without increased bandwidth requirements. A single channel carrier system providing an extra two-way speech circuit over an existing open wire line.

Compact low-priced companders to give improved performance on radio and carrier circuits suffering from excessive induced noise and crosstalk.

A compact 7-channel stackable rural carrier system suitable for bothway junction circuits and at a price which makes it economical to use for individual subscriber services.

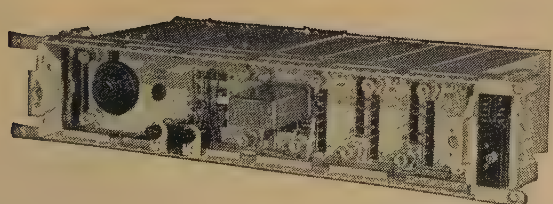


FREQUENCY DIVISION MULTIPLEX

TCS.12 equipment provides groups of 12 channels of C.C.I.T.T. performance in the 60—108 kc/s spectrum and assembled into groups and supergroups for frequency division multiplex systems.

Sub-groups of 3-channels are used and a feature of the equipment is the ease with which a programme circuit can be substituted for one of the sub-groups. Speech-immune signalling paths at 3825 c/s equivalent are employed and the signalling circuits will handle 17 c/s ringing or automatic dialling. Current requirements are approximately 1.5 amps at 24-50 volts D.C. per nine foot rack of 36 channels.

cuts cost space and maintenance



The associated frequency generating apparatus is controlled by a master oscillator (shown left) of high stability. Duplicate equipment is supplied and this together with automatic changeover provides a service of extreme reliability. Frequency checking, monitoring and alarm facilities are incorporated in the equipment and the whole carrier supply apparatus can be readily extended to meet all carrier supply requirements from 12 channels upwards.

T E L

E116A

ERICSSON TELEPHONES LIMITED — ETELCO LIMITED

HEAD OFFICE: 22 LINCOLNS INN FIELDS, LONDON, W.C.2. TEL. HOL. 6936

WORKS— BEESTON, NOTTINGHAM AND SUNDERLAND.

vast experience in the science of telecommunications

As a result of years of experience in the world of telecommunications, we have an unsurpassed wealth of knowledge of this science, which is freely offered to you as part of the incomparable Siemens Ediswan service.

- Comprehensive Telecommunication Systems
- Public Telephone Exchanges
- Private Telephone Systems
- Telephone Instruments and Apparatus
- Carrier Telephone Systems
- F.M.V.F. Telegraph Equipment
- V.H.F. Radio Telephone Terminal
- Marine Signalling Apparatus
- Radio Equipment
- Control & Indicating Apparatus

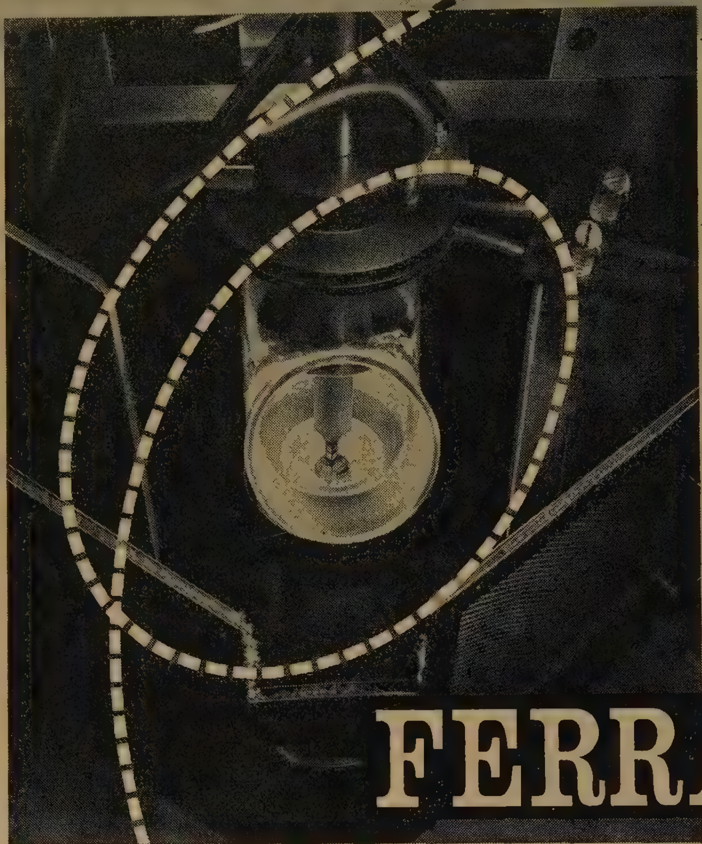
SIEMENS EDISON SWAN LTD.

An A.E.I. Company

Telecommunications Division, Woolwich, London S.E.18

Telephone: Woolwich 2020

The heart of the matter... the art of the matter



The formation of a single silicon crystal ingot

FERRANTI

offer the widest range of

SILICON Semiconductor Devices

in the United Kingdom

Ferranti Ltd. were the first company in Britain to introduce Silicon semiconductor devices as used in magnetic amplifiers, in aircraft, guided missiles, radar and computers. Until recently they were the only firm in the United Kingdom supplying silicon diodes in quantity. Commencing their programme of research and development in 1954, they have already made outstanding contributions to technique, and are now producing at Gem Mill, Oldham well over half-a-million silicon diodes annually in the widest range offered by any British manufacturer.

Among the 120 or more different devices are rectifiers, fast diodes, zener diodes, tetra-layer diodes and triodes, alloy junction transistors, diffused junction transistors, photo-voltaic cells, voltage variable capacitors and many new ones.

Data Sheets, Application Reports etc., advice and assistance in techniques of application are freely available.

FERRANTI

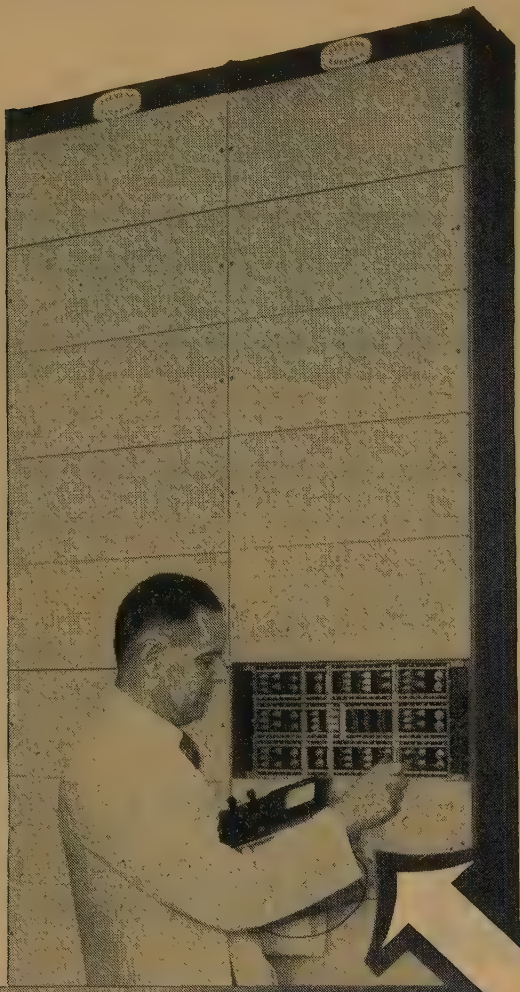
SILICON SEMICONDUCTOR DEVICES

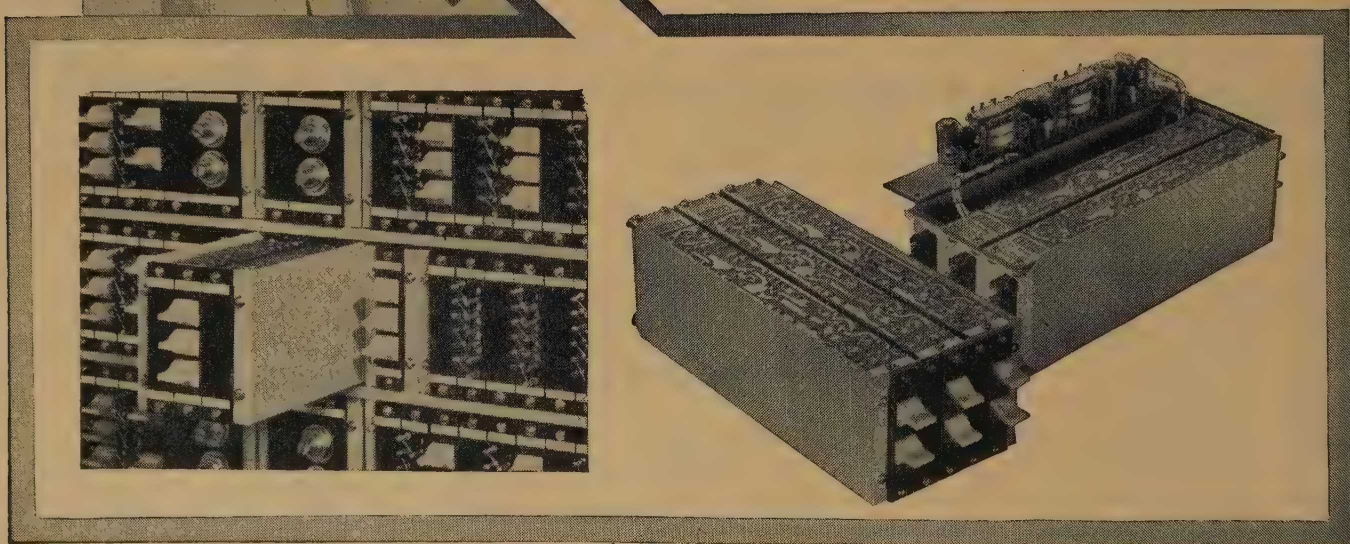
FERRANTI

FERRANTI LTD · GEM MILL · CHADDERTON · OLDHAM · LANCS · *Telephone:* MAIn 6661

London Office: KERN HOUSE · 36 KINGSWAY · W.C.2 *Telephone:* TEMple Bar 6666

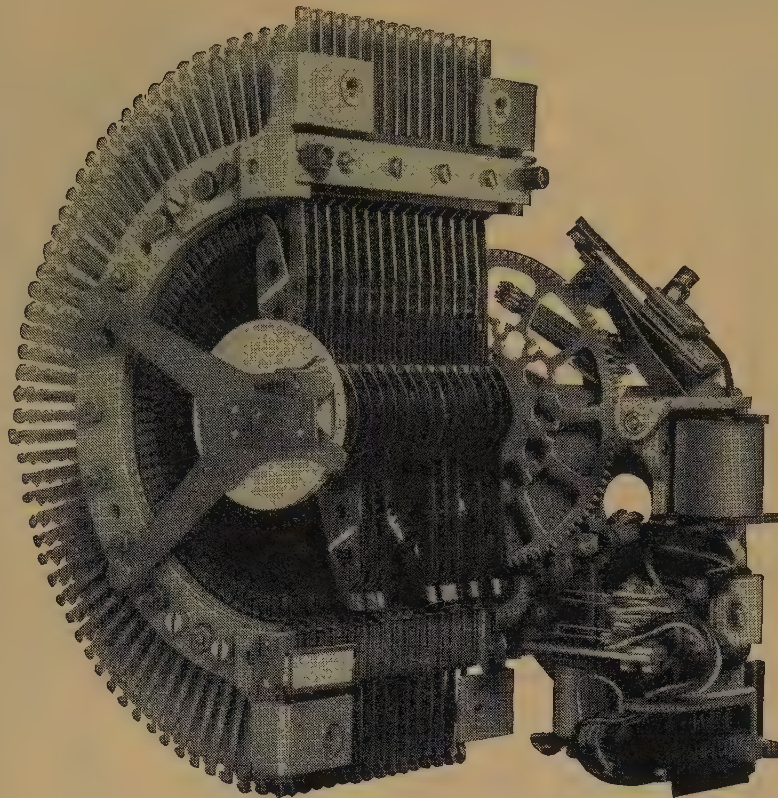
FE 190





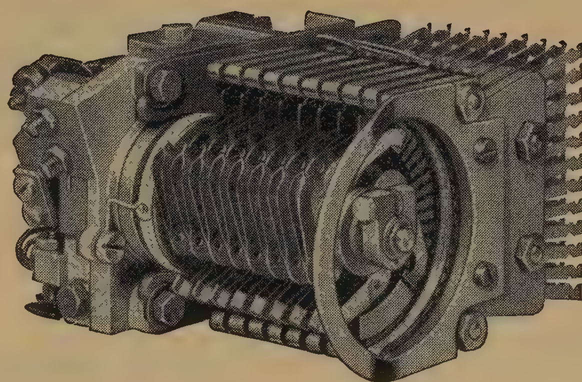
These 3 outstanding uniselectors

provide long life and reliability in operation



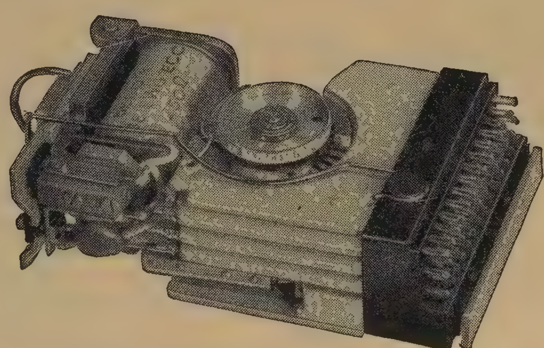
The high speed motor uniselector

has an availability of 52 outlets in a half revolution with up to 16 parallel channels, and is capable of high speed search if required. It is also suitable for scanning and other collection functions, or for the routing of information to a multitude of points. It can be brought to rest at any desired outlet by means of simple circuitry, and a homing drive is provided so that it can be readily restored to the normal position. Arrangements can also be made for step-by-step working.



The digit switch

is specifically designed for step-by-step operation in response to impulses, and has an availability of 12 outlets with 8 parallel channels. It may be used to set up a marking condition for the control of the Motor Uniselector.



The miniature uniselector

is a compact switch of unique design which occupies no more space than a standard 3000-type relay. It has an availability of 12 outlets with 3 parallel channels. A great advantage is that it incorporates a plug-in type of mounting so that the entire device can be removed and replaced without detaching any soldered connections.

SIEMENS EDISWAN

Full technical information will gladly be supplied on application

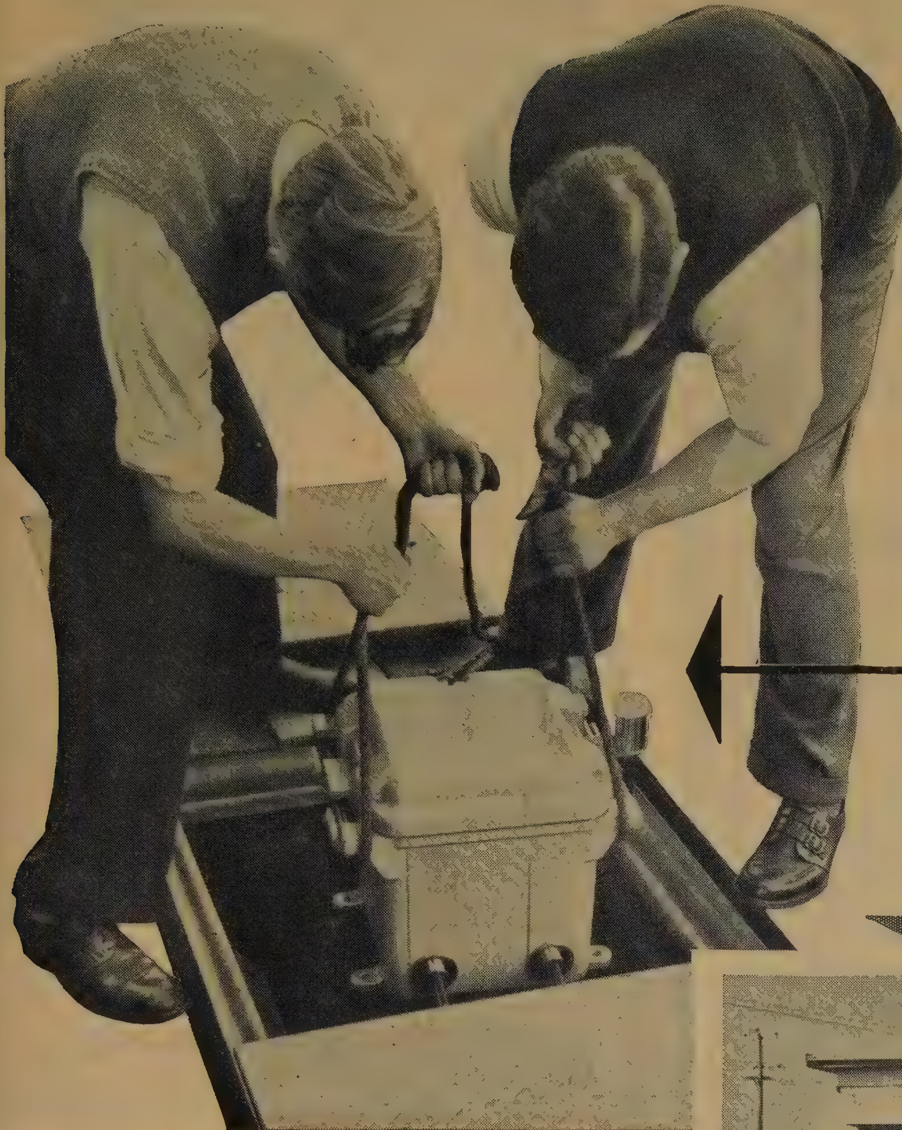
SIEMENS EDISON SWAN LTD

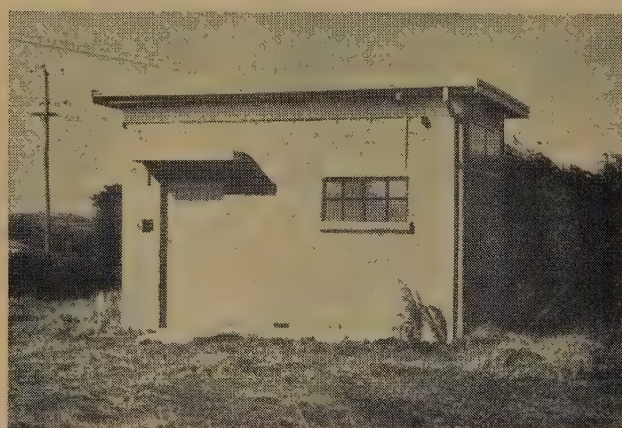
An A.E.I. Company

Telecommunications Division, P.D.7.
Woolwich, London S.E.18.
Telephone Woolwich 2020 Ex. 326

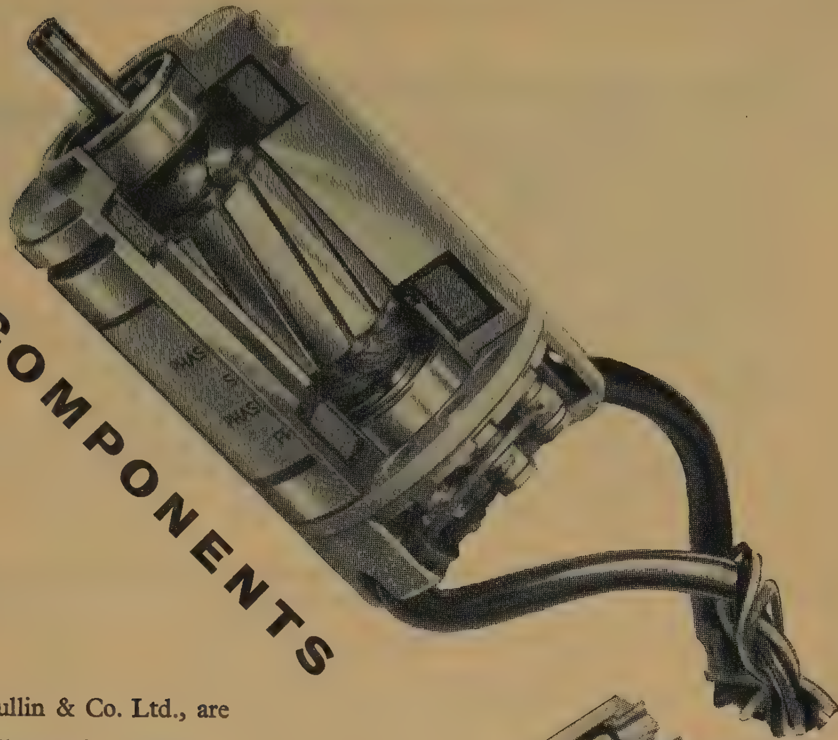
PD 7/3

TA 2139

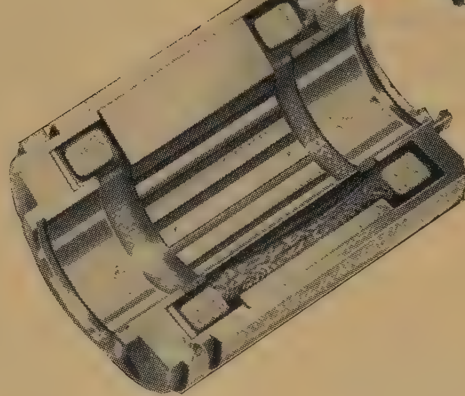




Araldite IN SERVOCOMPONENTS



The Synchro units shown, made by R. B. Pullin & Co. Ltd., are sectioned to show how the stators are integrally cast in Araldite to provide maximum protection against the effects of extremes of temperature, humidity and vibration. The excellent machining properties of Araldite make possible a straight-through bore technique, which eliminates errors in alignment and also permits the use of the smallest possible air gap between rotor and stator. High insulation and dielectric strength, remarkable adhesion to metals, and negligible shrinkage on curing make Araldite eminently suitable for use in the construction of precision electrical equipment.



Araldite epoxy resins are used—

- **for casting high grade solid electrical insulation**
- **for impregnating, potting or sealing electrical windings and components**
- **for producing glass fibre laminates**
- **for producing patterns, models, jigs and tools**
- **as fillers for sheet metal work**
- **as protective coatings for metal, wood and ceramic surfaces**
- **for bonding metals, ceramics, etc.**

This photograph shows an A.E.W. electric oven, capable of maintaining temperatures within close limits, as used by R. B. Pullin & Co. Ltd. for curing the Araldite-filled stators.

Araldite *epoxy resins*

Araldite is a registered trade name

CIBA (A.R.L.) LIMITED

Duxford, Cambridge   Telephone: Sawston 2121

AP 463

Semiconductors

COMPUTER TRANSISTORS

The Semiconductors range of Computer Transistors, designed and tested to the special requirements of computer engineers, is the key to a new order of computer speed and reliability. Overall reliability is further increased by making possible a substantial reduction in the number of associated components.

The two types of Silicon Alloy Transistor shortly going into production will make it possible to extend this high-speed computer performance into **ambient** temperatures well above 100°C. **Samples are available now.**

	TYPE	DESCRIPTION	RISE TIME millimicroseconds	Vc max	Ic max
HIGH-SPEED LOW-LEVEL SWITCHING **GERMANIUM**	SB 344 SB 345	General purpose transistors for conventional logic circuits.	50	5v	5mA
	SB 240	Designed for directly coupled circuits. Controlled input, saturation and hole storage characteristics.	30	6v	15mA
	MA 393	High gain transistor for high-speed driving of parallel circuits.	30	6v	50mA
	2N 501	Ultra-high speed transistor with controlled input and saturation characteristics.	10	12v	50mA
HIGH-SPEED LOW-LEVEL SWITCHING **SILICON**	SA 495	General purpose 10Mc/s transistor for conventional logic circuits.	100	25v	50mA
	SA 496	15Mc/s transistor for directly coupled circuits. Saturation resistance typically 10 ohms. Controlled input and hole storage characteristics.	80	10v	50mA
CORE DRIVING **GERMANIUM**	2 N 597 2 N 598 2 N 599	min $f\alpha$ 3Mc/s min $f\alpha$ 5Mc/s min $f\alpha$ 12Mc/s 250 mW high frequency alloy transistors with high gain and low saturation resistance	400 * 250 * 100 *	20v 20v 20v	400mA 400mA 400mA
	2 N 600 2 N 601	min $f\alpha$ 5Mc/s min $f\alpha$ 12Mc/s 750 mW versions of 2 N 598 and 2 N 599. Peak current 3 amps.	250 * 100 *	20v 20v	400mA 400mA

* *rise time to 400mA*

Full technical details and applications assistance available on request.

Semiconductors Limited

CHENEY MANOR
SWINDON · WILTS
TELEPHONE: SWINDON 6421/7

S.C.I

INDEX OF ADVERTISERS

SILICON RECTIFIERS

up to 600 Volts P.I.V.

3 amps

TEXAS SILICON RECTIFIERS offer excellent high temperature characteristics. For example the stud-mounted type gives an output of 3 amperes at 50°C and will still provide 1 ampere at 150°C with no change in the rated peak inverse voltage. The rugged, welded housing with glass-to-metal seal provides high resistance to shock and vibration. Listed in the table below are the characteristics of representative devices of each type.

Have you received your copy of the latest Texas Application Report "D.C. Power Supply Circuits using Silicon Rectifiers"? If not, or if you require fuller details of Texas Rectifiers, please write your name and address in the margin and return this advertisement to us.

	Symbol	1S001	1S005	1S111	1S115	1S401 1S401R	1S405 1S405R
MAXIMUM RATINGS							
Peak Inverse Voltage at −65°C to +150°C	PIV	200V	600V	200V	600V	200V	600V
Average Rectified Forward Current at +50°C	I_o	750mA	750mA	† 400mA	† 400mA	*3A	*3A
Average Rectified Forward Current at +150°C	I_o	250mA	250mA	150mA	150mA	*1A	*1A
Recurrent Peak Forward Current at +50°C	i_l	† 2·5A	† 2·5A	† 1·25A	† 1·25A	*10A	*10A
Surge Current for 10 Milliseconds	I_{PK}	16A	16A	6A	6A	33A	33A
Operating Temperature, Ambient	T_A	—65°C to +150°C					
SPECIFICATIONS							
Minimum Breakdown Voltage at +150°C	V_z	240V	720V	240V	720V	240V	720V
Maximum Reverse Current at P.I.V. at +25°C	LI_b	10μA	10μA	0·2μA	0·2μA	10μA	10μA
Maximum Forward Voltage Drop at +25°C	E_b	1·0V	1·0V	1·0V	1·0V	1·1V	1·1V
		(I_o=500mA)		(I_o=400mA)		(I_b=1Amp)	

* Rectifier mounted on 2″ x 2″ x $\frac{1}{16}$″ aluminium Heat Sink † @ 25°C

TEXAS INSTRUMENTS LIMITED

TELEPHONE: BEDFORD 68051 DALLAS ROAD BEDFORD CABLES: TEXINLIM BEDFORD

THE PROCEEDINGS OF
THE INSTITUTION OF ELECTRICAL ENGINEERS

EDITED UNDER THE SUPERINTENDENCE OF W. K. BRASHER, C.B.E., M.A., M.I.E.E., SECRETARY

VOL. 106. PART B. NO. 29. SEPTEMBER 1959

621.317: 550.3

The Institution of Electrical Engineers
Paper No. 3035
Sept. 1959
©

THE FIFTIETH KELVIN LECTURE

'THE INTERNATIONAL GEOPHYSICAL YEAR 1957–58'

By Sir DAVID BRUNT, Sc.D., F.R.S.

(*Lecture delivered before* THE INSTITUTION, 23*rd April*, 1959.)

The International Geophysical Year, 1957–58, has been the most astounding feat of international co-operation in the whole history of the human race. In its earliest initiation it was intended as a continuation of the First International Polar Year of August, 1882, to September, 1883, and of the Second International Polar Year in 1932–33, both of which were limited to high latitudes. When the suggestion of a third International Year was first made in 1952, geophysicists had already become familiar with the nature of the ionosphere, and how it was greatly influenced by solar activity, as are geomagnetism, the aurora, and cosmic rays, which are strongly influenced by the earth's magnetic field. The whole field of geophysics had been widened as knowledge increased, and instrumental equipment, notably that used for observing the ionosphere, had been enormously improved since the Second International Polar Year. Automatic instruments were by this time available, and it seemed highly desirable that the new International Year should cover the whole field of geophysics, and make full use of the improved techniques.

After very careful consideration, this idea was universally accepted, and it was agreed that the field of observation during the Geophysical Year should cover (*a*) all the phenomena which are directly affected by solar activity as shown by sunspots and bright 'flares' on the sun's surface, and (*b*) other phenomena of which our knowledge would be strikingly increased by observation at a widespread network of stations, and in particular by observations at new stations in regions where hitherto few or no observations had been made. The time covered by the Geophysical Year was selected to fall at the time of maximum sunspot activity, 1957–58. The clearing house for sunspot statistics, the Zürich Observatory, later reported that 1957 actually had the highest yearly mean of relative sunspot numbers since the initiation of sunspot records in 1778. The all-time record of sunspot activity occurred about half-way through the I.G.Y. period. Thus the period chosen for the I.G.Y. turned out to be miraculously fortunate.

The programme may be regarded as falling into three broad categories:

(i) *Phenomena in the atmospheres of earth and sun.*—Solar activity, meteorology, the ionosphere, geomagnetism, aurora and airglow, cosmic rays.

(ii) *Investigations of the earth's surface.*—Oceanography, glaciology, precise measures of latitude and longitude.

(iii) *Investigation of the earth's interior.*—Earthquakes, gravity measurements.

The following table shows the number of stations operating in each of the 13 disciplines represented in the observational programme.

Subject	Number of stations (at 31st May, 1957)	Number of United Kingdom stations (some jointly with other nations)
Meteorology	1 603*	65*
Geomagnetism ..	276	10
Aurora and air glow ..	270*	18*
Ionosphere	291	25
Solar activity	119	5
Cosmic rays	136	5
Longitude and latitude	78	3
Glaciology	84	7
Oceanography	200	10 + 5 ships
Rockets and satellites ..	47	1
Seismology	335	11
Gravity measurements	191	6
Nuclear radiation ..	—	20
	3 630	191

* These are principal stations only; there are many auxiliaries.

It should be noted that a station is here to be defined as a site where observations in one discipline are made. A site where two disciplines are observed counts as two stations, and so on.

The table takes no account of the many subsidiary and 'voluntary' sites at which observers have co-operated, and allowing for the fact that all published lists are likely to have become incomplete, we can safely say that a total of 4 000 active sites is almost certainly an underestimate. It is probable that some 30 000 scientists, engineers and technicians were engaged in the effort, while the volunteer observers may have added another 20 000 to the number of workers. The total cost can only be guessed at, though a figure of £500 million per annum has been suggested. In 1952, when scientists began to formulate their programmes of observation, to improve their instruments, and to devise entirely new instruments capable of measuring what had not hitherto been measured, there was no expectation of the rapid development of rockets or of artificial satellites. When rockets

Sir David Brunt is Chairman of the Electricity Supply Research Council.

and satellites became available to carry instruments to great heights, they provided a powerful new tool for geophysicsts, capable of observing and recording the conditions up to great heights in the atmosphere. With the use of these new tools a new era began in the study of the ionosphere, and of space up to considerable heights above the earth's surface.

So vast is the number of observations accumulated during the I.G.Y. that it will need intensive work by very many teams of workers to reduce all these observations and to deduce their interpretation. At the present time, while we can definitely state that some quite startling new facts have already emerged, it is certain that the full interpretation of the main body of the observations will not be achieved for some years. Consequently, it must be admitted that some of the statements which I shall make may later need modification. No complete or final appraisal of the results of the achievements of the I.G.Y. can yet be expected. All that can be attempted here is to draw attention to the outstanding new concepts which have already been established, or alternatively appear to have a considerable probability of being true.

SOLAR ACTIVITY

Before we proceed to discuss any observations, a brief account must be given of arrangements made to inform observers of the likelihood of disturbances on the sun's surface.

It is not possible to foretell, more than a few days ahead, the days of notable intensification of radiations from the sun, so that a time-table of such intensifications cannot be produced well in advance. From a close study of the sun, however, it is generally possible to say a few days in advance whether such a disturbance will probably occur. A World Alert will then be announced, so that all observing stations will be ready to make specially intensive observations. On subsequent days all available fresh information is considered, with a view to cancelling the alert if it appears that the expected disturbance is no longer likely to develop.

This has required the organization of a very close watch of the sun for signs of solar activity. Some 50 stations have made regular observations, one half of them using spectrohelioscopes, the remainder being equipped with automatic helioscopes, which photograph the sun's disc through Lyot-type filters, at intervals of either one or three minutes. The watch on the sun was fully maintained through 89·7% of the possible hours during 1958.

A notable British contribution to this programme has been made at the Royal Observatory at the Cape of Good Hope, where a Lyot Hα heliograph was installed early in 1958, on the recommendation of the British I.G.Y. Committee. This instrument recorded over 600 flares in its first nine months of operation.

In each month, three days were selected as Regular World Days, on which specially intensive observations in all disciplines should be made at all the world-wide network of stations. The three days were fixed, one on the day of new moon and the two others on days associated with unusual meteoric activity. For the purposes of meteorology a period of ten consecutive days, to be known as a World Meteorological Interval, was fixed in each quarter, coming at the winter and summer solstices, and at the two equinoxes.

COLLECTION OF OBSERVED DATA AT WORLD DATA CENTRES

The observational material resulting from the programme of observations is being collected at three World Data Centres (W.D.C.), each of which will eventually have a complete set of all observations.

World Data Centres

Centre A, U.S.A.	Centre B, U.S.S.R.	Centre C
With 11 sub-centres at different institutions (all subjects)	3 sub-centres in Moscow, 1 in the Crimea (all subjects)	Meteorology in W.M.O. Switzerlan Geomagnetism (Denmark, Japan) Aurora (United Kingdom, Swede Airglow (France, Japan) Ionosphere (United Kingdom, Japan) Solar activity (Switzerland, Ital France, Germany, Australia) Cosmic rays (Sweden, Japan) Glaciology (United Kingdom) Rockets and satellites (Unite Kingdom) Seismology (France) Gravity (Belgium) Nuclear Radiation (Sweden, Japa

The World Data Centres in the United Kingdom are locate as follows:

Aurora.—Balfour Stewart Auroral Laboratory Department o Natural Philosophy, University of Edinburgh.

Ionosphere and rockets and satellites.—D.S.I.R. Radi Research Station, Slough.

Glaciology.—British Glaciological Society, housed at the Sco Polar Research Institute of the University of Cambridge.

The W.D.C. for Meteorology will be cared for by the Worl Meteorological Organization (W.M.O.) in Geneva.

Each of the three World Data Centres will receive origina copies of I.G.Y. data from countries that elect to use that pa ticular centre as the archives for their data, and will then supp copies to the two other W.D.C.'s. Some idea of the magnitud of the data-collecting operation may be gathered from the fa that eventually the W.M.O. will have in its archives ove 100 million meteorological observations. The W.D.C. for th ionosphere, at Slough, has already received some 800 000 photo graphic records, graphs and tables of data. These two example give some idea of the magnitude of the task of those who wi reduce and interpret the I.G.Y. data.

EMISSIONS FROM SOLAR FLARES: IONOSPHERIC EFFECTS

A solar flare is a sudden blazing up, to about ten times i normal brilliance, of an area of the troposphere of some hundre of millions of square miles. Solar flares are not visible to th eye, but show up very brightly against the normal backgroun of the troposphere in hydrogen light. They are best observe by the light of the hydrogen line with a spectrohelioscope, an are found in close association with sunspots. A magneticall complex sunspot may produce 30–40 flares during a sing passage across the sun's disc.

Flares are classified on a visual scale of importance, fro Class I (smallest) to Class 3 (largest), with Class 3+ to indica flares of exceptional area and intensity. The average duratio of Class I is about 17 min, of Class 3 about 1 h, and of Class 3 about 3 h, and the area of the flare can range from 100 to we over 1 000 in millionths of the area of the sun's hemisphe (Fig. 1).*

The flares emit (*a*) wave radiations, which travel with th speed of light and reach the earth in 8·3 min; magnetic stor particles, which are ions and electrons and travel much mo slowly, taking 20–40 h to reach the earth; and (*c*) cosmic-r particles, mainly protons, which travel with an intermedia speed, taking up to nearly an hour to reach the earth.

The wave radiations are of three classes, (i) ultra-violet a

* A millionth of the sun's hemisphere has an area of 3×10^6 km^2, or 1·17 10^6 sq miles.

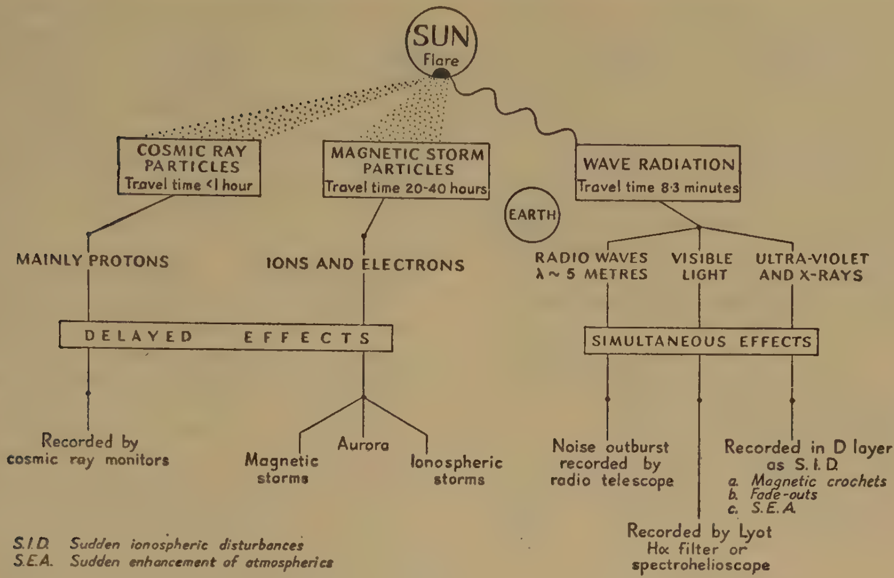


Fig. 1

Reproduced through the courtesy of Professor M. A. Ellison, of the Dublin Institute of Advanced Studies.

X-rays, which penetrate downwards as far as the D-layer, the lowest layer of the ionosphere, where X-rays in the waveband 1–8 Å produce additional ionization, yielding an increased supply of free electrons and a sudden increase of the electric current flowing in the D-layer, giving a kink, or crochet, in the records of magnetic storms. The extra ionization increases the absorption of short radio waves on their way down from the higher F_2-layer, producing a fade-out of radio signals. The fade-out begins suddenly during the flash of the flare, as it rises to maximum brightness, which may be attained in less than one minute. If the path of the radio signals passes near the sub-solar point, i.e. the region where the sun is overhead at the time, the signal strength drops to $\frac{1}{5}$–$\frac{1}{10}$ of its original value. During the flare, the D-layer has enhanced power of reflection for a distant transmitter which gives a sudden enhancement of atmospherics; the ceiling falls, the concentration of electrons necessary for reflection then being found at a lower level. Since the waves emitted from the flare all travel with the same speed as light, their effects are first seen at the same instant as the flare is seen. Thus the time of the flashing up of the flare can be recorded by the magnetic crochets or the sudden enhancement of atmospherics.

(ii) The radio waves from a flare can be observed by radio telescopes, and in England such observations are made at Jodrell Bank and at Cambridge. A sudden outburst of noise recorded by a radio telescope will also give the time of initiation of the flare.

(iii) The visible light from the flare is recorded by a Lyot Hα filter or a spectroheliosope.

The particles emitted by the flare have a wide range of speeds. The slower-moving particles, which reach the earth some 26 h after the waves, with an average speed of 1 000 miles/sec, are probably ionized atoms and electrons, which on arrival at the earth yield great currents of electricity, which produce great magnetic storms, and aurora in the polar regions, and at times aurorae extending to the equator.

It has been found that a magnetic storm is likelier to follow a flare seen near the centre of the sun's disc; when the flare is near the limb of the sun, the earth is less likely to be in the line of particle fire. From this it is concluded that the particles are emitted from the sun in a vertical direction. Parts of the sun's disc appear to be regions of continual emission of particles, of the nature of jets which rotate with the sun, and strike the earth at intervals of about 27 days.

The fast-moving particles emitted by flares are believed to be mainly protons, with exceptionally high energies, varying from 10^8 to 10^{17} eV, travelling with speeds of an appreciable fraction of the speed of light, and capable of disintegrating atmospheric atoms. The slow-moving particles are deflected towards the regions of the two geomagnetic poles, and only the particles of high energy penetrate into middle latitudes.

It is to be noted that the phenomena due to ultra-violet and X-rays, radio waves and visible light have their initiation at the time of the first flash of the flare; the magnetic storm and the arurora appear about one day later.

United States rocket observations made during the total eclipse of the 12th October, 1958, showed the existence of a powerful flare of X-rays in the waveband 1–8 Å. Four rockets instrumented to measure the intensity of ultra-violet and X-radiation showed that the ultra-violet diminished to zero as the stage of totality was reached, while the X-ray intensity remained substantially unchanged, showing that the ultra-violet radiation emanates from the sun's disc, or chromosphere, while the source of the X-radiation is the sun's outer atmosphere, or corona. But it has not yet been established whether the sources of the ultra-violet and X-radiation are localized in special regions of the chromosphere and the corona, respectively.

The question, 'What is a flare?' is one to which no answer can be given. A flare covers a region of the chromosphere several hundred million square miles in extent. It increases its intensity of radiation by a factor of at least ten, and achieves its maximum intensity in many cases in less than one minute. I am not tempted to speculate on the nature of the mechanism that works so rapidly and so efficiently.

The detailed study of the magnetic fields in active solar regions carried out in the United States and the Soviet Union have shown that at the time of a flare the lines of force in the adjacent field undergo a sudden redistribution, suggesting that in the flare there is a sudden conversion of magnetic energy into wave and particle radiation. This may be the key to the secret of the flare.

A major solar flare was observed on the 28th June, 1957, and

charged particles from this flare reached the earth two days later, producing a severe magnetic storm. There were severe disturbances in the ionosphere, and long-range radiocommunication was blacked out for a considerable time. An Alert was issued at 1600 h on the 28th June, and a Special World Interval was declared on the 29th June for 30th June, 0001 h. Pulses transmitted vertically to the ionosphere were not reflected but were absorbed. On the 4th July it was found that the normal ion distribution above the D-layer remained undisturbed during the black-out, but for some 12 miles below the D-layer there was markedly increased ionization, later shown to be due to the X-ray emission from the sun, in the range of 1–2 Å and associated with solar flares. It is seen that the times of occurrence of the different phenomena fit in with the system described earlier.

One of the most striking observations of conditions in the very high atmosphere was obtained by Dr. J. A. van Allen and colleagues from observations made by instruments in Explorer I, which showed that, up to a height of 700 km, the intensity of cosmic rays was in accordance with expectation, but the instrumental behaviour above this level pointed to great increase in the intensity of radiation, which was interpreted as corpuscular in nature. Later satellites—Explorers III and IV—were specially instrumented for further investigation of this radiation, and later this research was continued in Pioneer III on the 6th December, 1958. It was found that there are two distinct, widely separated zones of high intensity of radiation. Van Allen and his colleagues believe it is now established that the great radiation belt around the earth consists of charged particles, temporarily trapped in the earth's magnetic field. The van Allen radiation reaches its peak intensity at a height of two earth radii, and extends up to eight earth radii.

Other methods of determining the constitution of the atmosphere at very high levels are being developed. Thus at very low radio frequencies, whistle-like sounds are detected. These are thought to originate in lightning flashes at the surface of the earth, and to travel thousands of miles along the earth's lines of magnetic force before returning to the opposite polar hemisphere. Recent observations suggest that the ion density and molecular density along the whistler path at altitudes of as much as twice the earth's radius are much greater than hitherto anticipated. It appears that the atmosphere extends far beyond the level at which it was supposed to end, and that a tenuous atmosphere—the sun's corona—fills all space between the earth and the sun.

By means of a radio-frequency mass spectrometer, it has been possible to separate the components of ionized gases between 230 and 930 km. Ions of atomic oxygen predominate, but ions of atomic nitrogen are also present, equivalent to 3–7% of the oxygen. There appears to be detectable number of ions in the region of 1 000 km, hitherto regarded as empty space. At 266 km, the density has been estimated to be 3×10^{-11} g/m^3, or approximately 10^{-10} times the surface density. At about 370 km, the density is about one-tenth of this value.

Sputnik III yielded measurements of 500 000 positive ions per cubic centimetre and 160 000 per cubic centimetre at heights of 242 and 795 km, respectively.

Two Soviet scientists, writing in *Pravda*, have discussed the inferences to be drawn about the density of the upper atmosphere from Sputnik observations. They state that at orbital heights the concentration of gas is greater, and the density of micro-meteorites considerably less, than had been anticipated. The micro-meteorites are not solid particles of stone or metal, but resemble flakes of snow or soot, and have a density of only a fraction of that of water. The densities of the atmosphere deduced from the braking effect upon the Sputniks at perigee were substantially greater than could have been expected on the basis of any information previously available.

The values deduced by these various methods are all in reason able agreement, and all yield values for the densities at grea heights which are substantially greater than have hitherto bee accepted. In high latitudes, rockets launched in winter an summer, by day and by night, indicate strong solar control o density in the high atmosphere, with a latitude effect, a seasona effect and a strong diurnal effect, none of which appear in lowe latitudes. The gases in the atmosphere appear to be well mixe up to 100 km, above which gravity separation begins to exert a effect.

The observations made by the use of rockets and satellite appear to confirm the theory that the terrestrial atmospher extends far beyond the level where it had previously been regarde as 'ending'. The view that the earth is embedded within th sun's corona is perhaps the most surprising and novel resul hitherto derived from the I.G.Y.

The United Kingdom programme of ionosphere observatio has been carried out as planned, and some very interesting nev information has been obtained. At Halley Bay, the Roya Society base in Antarctica, it has been found that the maximur ionization density in the F_2-layer at noon in winter exceeds th same quantity at noon in summer, in spite of the fact that at th height of the F_2-layer the sun does not rise in winter or set i midsummer. Also, there is a strong diurnal variation in winte giving electron densities at noon which are ten times those a midnight. The diurnal variation is small in summer, givin maximum density at noon less than at midnight. Near th equinoxes there is a sudden change from typical winter to typica summer diurnal variations, or vice versa. From these observa tions it is concluded that the variations of ionization density ar mainly due to movements in the ionosphere, and are only ver slightly affected by the direct ionizing action of the sun.

METEOROLOGY

The observations in meteorology are being collected by th World Meteorological Organization (W.M.O.) at Geneva. Thes will include those surface observations obtained from probabl some 2 100 stations, observations of winds and temperatures i the free air, of atmospheric ozone, of solar and atmospheri radiation, and radiation balance. The main object of th intensified meteorological programme is to provide fuller know ledge of the general circulation of the atmosphere, and to ascer tain the mechanism by which air masses are exchanged betwee tropical and polar regions. With this in view, a network c meteorological stations, covering the antarctic region, has bee created and is contributing to the W.M.O. a stream of observa tions which will be of great assistance in solving the mai problem. Some writers have expressed great confidence in th possibility of the study of these observations, and those of th distribution of ice in the polar regions, leading to improvemen in weather forecasting. How far it will be possible to realiz these hopes will only be known after a long period of reductio and study of the great mass of observations, not only those mad during the I.G.Y., but also those to be made during later yea of co-operation. The main synoptic attack on the meteor logical problem is yet to come.

The study of observations of weather in the Antarctic h already yielded interesting results. The antarctic ice-balance maintained by the intrusion of warm air-masses into the centr ice-cap region, and not, as hitherto supposed, by a strong hig altitude cyclone more or less permanently situated above th region. There is no permanent wind barrier round the continen as thought hitherto. Antarctic weather therefore plays an acti part in the atmospheric circulation of the southern hemispher There are six stable high-pressure tongues, extending northwa

in association with those parts of Antarctica that jut out from the main continental mass. They block the movement of cyclones, thus trapping them in relatively stable positions. The six cyclonic zones between the pressure ridges are over the Weddell Sea, off the eastern section of Queen Maud Land; over the Mackenzie Sea, off Wilkes and King George V Land; and over the Ross Sea. The snow which perpetuates the glaciation comes from cyclones which periodically intrude into a rim zone, 300–400 miles wide, round the area. Ridges of the antarctic anticyclones may join those in the sub-tropical regions and inhibit zonal circulation. Dissemination of meteorological observations by the Antarctic Weather Centre has already produced a significant improvement in the forecasts of southern hemisphere countries. The reduction of the I.G.Y. observations of upper air winds and temperatures may, however, well lead to a modification of views hitherto arrived at, with regard to the interpretations outlined above.

The use of rockets for atmospheric soundings has already led to a considerable increase of knowledge of the upper air, while the artificial satellites have shown that we can now probe into conditions at vastly greater heights. A very useful and ingenious electrolytic absorption apparatus devised by Dr. A. W. Brewer of Oxford, to measure the concentration of ozone, has been flown by sounding balloons up to about 20 km at Tromsö, Malta and Halley Bay.

A number of studies in certain special aspects of meteorology have been made by participating countries. Among these we may mention the measurement of the radioactive matter in the atmosphere, the CO_2 content, and the ozone content. An especially interesting study made at the U.S. drifting station *Alpha* in the Arctic Ocean, yielded the result that, during the melt season, melting proceeds much more rapidly under overcast sky than under clear sky with the sun shining. The systematic study of the general circulation of the atmosphere must await a later stage.

GEOMAGNETISM AND AURORA AND AIRGLOW

The major portion of the earth's magnetic field is thought to be generated by electric currents within the earth. The remaining part, about 5% of the total, is generated externally, probably by the great earth-girdling systems of electric currents in the charged, or ionized, regions of the upper atmosphere. The internally generated field is relatively stable and undergoes slow, so-called 'secular', changes measured in periods of possibly years or centuries. The externally generated field is marked by rapid, transient fluctuations, measured in periods which may be days, hours or even seconds.

Geomagnetic observations have been made at four stations in the United Kingdom, at Argentine Island and at the station at Halley Bay in Antarctica. These observations, and those made elsewhere by other nations, will be examined in conjunction with observations of aurora, of conditions in the ionosphere, and of solar activity. The geomagnetic observations at Halley Bay will be of special value in this comparison, since Halley Bay is just outside the zone of greatest concentration of the ionospheric electric currents, which cause geomagnetic disturbances, and which are also associated with the greatest frequency of aurora. It will now for the first time be possible, since a world-wide network of stations covering all latitudes has been set up, to study the geographical distribution of aurora, as well as the details of changes in form and movement during a great display of aurora. Observers on land, on ships at sea, in aeroplanes, and on expeditions to the polar regions, are all making a contribution to the recording of aurora, while in higher latitudes all-sky cameras have continually photographed the whole sky. Radio-echo methods are used to detect daytime aurorae.

The I.G.Y. has been notably rich in aurorae. An interesting result already achieved by British observations is that simultaneous displays of aurora borealis and of aurora australis progress in a similar manner, while, near the solstices, aurora penetrates further into low latitudes in the summer than in the winter hemisphere. The zones of equal frequency of occurrence of aurorae are found to be more nearly along lines of constant magnetic inclination than along lines of geomagnetic latitude, as had hitherto been assumed.

In addition to the aurora, a faint atmospheric luminosity known as the 'airglow' is visible at night all over the world. Its distribution can now for the first time be discussed on a global scale, and it should be possible to determine whether it is closely related to the aurora.

LONGITUDES AND LATITUDES

This is a very long-term investigation, concerned with time and latitude variation, and although the work has been intensified slightly, it is not at present possible to assess its geodetic value. It must be remembered that we do not yet know the distances between the continents, with accuracy.

GLACIOLOGY

One-tenth of the earth's surface is covered by ice, providing a major source of cold air. During the I.G.Y. some 23 nations have made a close study of the ice cover, either of the smaller ice-masses such as those of glaciers in the Alps, the Rockies, the Altai, the Himalayas and the Andes, or the great polar ice-sheets of Greenland and Antarctica. Most of the mountain glaciers in Europe and North America have for some time been in retreat, as a result of the prevailing improvement of climate in the present century. It is found that the variations in the glaciers of South Georgia, at 54° S, agree with those in the North Atlantic, except for a lag of the order of 10–15 years. The culmination of the advance of the ice took place about 1891 in the north, and about 1900–1905 in South Georgia. A further examination is being made of the relation of variations of meteorological conditions with the amount of sea ice.

The equatorial tropical glaciers of Mount Kenya, Kilimanjaro and Ruwenzori, in East Africa, have all been under close observation, and it has been found that the glaciers of Kenya and Ruwenzori have diminished in size since they were first seen 60 years ago. It will be of interest to establish whether these fluctuations are representative of the fluctuations of rainfall of the region in the days before European settlement there.

Few of the results of the ice programmes are yet available, but it is known that there is a relationship between glacier behaviour and the small but significant fluctuations of climate over Europe and America in the past few centuries. It is thought that this marks the way to a solution of the problem of seasonal forecasting.

The ice cover of Antarctica has been the subject of very considerable investigation during the I.G.Y. The results are by no means easy to summarize, since the distribution of the ice over the deeply indented solid surface beneath it is extremely complex, and observations during traverses over the ice made by different nations have yielded widely varying depths. The thickest ice-cover ever measured, some 100 miles east of Byrd Station, is about 14 000 ft thick, and rests on a rock bed 8 200 ft below sea-level. Over a considerable part of Antarctica the rock floor is below sea-level.

The investigations in Antarctica have shown that the amount of ice over it is far greater than had been previously thought. As a result it is necessary to increase by 40% the estimate of the world's total of ice, to about 4·5 million cubic miles.

OCEANOGRAPHY

Preliminary study of the I.G.Y. recording of long waves suggests a relation to weather, including hurricanes, but the voluminous observational material has not yet been examined. The programme on ocean circulation has revealed some amazing new facts. Dr. Swallow's acoustic signalling float, which can be made to drift with the water at any desired depth, has made it possible to investigate in detail the variation of the water motion with depth. Three major counter-currents in the ocean have been found and measured. Thus below the Gulf Stream, at about 6 500 ft, there is a layer of water with little or no horizontal movement, and below this, at a depth of some 9 000 ft below the surface, is a counter-current in which the southward transport of water, at a rate of about 8 miles per day, is roughly equal to that in the opposite direction in the Gulf Stream. The bottom waters are rich in chemical nutrients, and where they upwell are found the great fishing banks, such as the Grand Bank off Newfoundland and the Peruvian fishing grounds.

British ships, working in co-operation with Norwegian and French research vessels, have revealed another counter-current some 400 miles west of Oporto and Lisbon, flowing at 1–2 miles per day, at a depth of about 4 600 ft.

The swift sub-surface equatorial current in the Pacific, discovered in 1952 by Dr. Townsend Cornwall, has been mapped by two U.S. ships and found to be nearly as extensive as the Gulf Stream. It flows eastwards along the equator for 3 500 miles from 140° W along a front 250 miles wide, at about 3 knots, about 100 ft to at least 1 000 ft below the surface, against the equatorial current of the Pacific. A third sub-surface current, 200 miles north of the equator, is probably greater than the Townsend Cornwall current.

An amazing discovery is a mineral-rich region in the south-east Pacific with extensive regions covered with a sludge laden with manganese and iron, and some cobalt mixed with copper. The value of these metals is estimated at $500 000 per square mile, and the economics of dredging up the sludge 'appear promising'.

The detailed examination of the ocean observations made during the I.G.Y. may yield results as remarkable as those mentioned above, since we may now regard the study of ocean currents as having been firmly put on a three-dimensional basis.

As engineers, you may think that you have no need to trouble yourselves to consider the currents in the ocean. You have during the last quarter of a century been forced to consider how the winds dissipate your chimney emissions, and it seems probable that, when you jettison your radioactive waste materials into the sea, you will have to give very special consideration to where they may come up again to the surface.

ROCKETS AND SATELLITES

Upper-air sounding rockets have been launched by Australia, Canada, France, Japan, the Soviet Union, the United Kingdom, and the United States. Valuable information concerning radiation at great heights has already been obtained, and the Soviet Union is rendering further valuable service in launching rockets in Antarctica. Observations by rocket are accumulating, and their final discussion will be awaited with interest. The artificial satellites have pushed the zone of observation to still greater heights, and have given us a vastly changed outlook upon the relation of the earth's atmosphere to the sun. One of the most striking observations yet recorded is the discovery of the Van Allen radiation, already mentioned.

Another striking illustration of the unique potentialities of rockets and satellites is given by the observations of sources of ultra-violet radiation in the galaxy made by rockets. Ultra-violet emission from the high atmosphere cannot be observed from the ground, as it is readily absorbed in the lower atmosphere. It can, however, be readily observed from rockets or satellites. Aerobee rockets have located strong sources of ultra-violet radiation, which do not correlate either with naked-eye objects in the sky or with the radio stars. For example, around α-Virginis, extensive nebular sources of ultra-violet radiation have been observed, radiating in the band 1 225–1 350 Å, but there is no visible nebula in this region. In Orion there is also extensive ultra-violet radiation from areas not observable in visible light.

Most of the ultra-violet radiation comes from near the galactic plane, but strong sources have also been located in high galactic latitudes. It thus appears that to visual astronomy and radio astronomy we should now add ultra-violet astronomy.

The artificial satellites launched by the Soviet Union and the United States have been, whenever possible, observed in the United Kingdom, and British scientists have been able to deduce new information about the structure of the ionosphere from observations of the radio transmission of the satellites. Optical observations by kine-theodolites at Ministry of Supply Establishments have received international recognition as among the most accurate of their kind. From these observations valuable conclusions are reached affecting our knowledge of the frictional drag of the atmosphere.

The constructing and launching of satellites cost very large sums of money. For example, the U.S. Army's satellite Explorer I cost £1 400 000, or £140 000 for each pound of weight put into orbit. Only about 3% of the information radioed by the satellite was recorded.

SEISMOLOGY

The main effort of seismologists has been directed towards obtaining comprehensive data of such earthquakes as occurred during the I.G.Y. New stations, some to be temporary, some to be permanent, have been established in parts of the world where such stations were widely scattered or even non-existent. It is as yet too early to look for a comprehensive discussion of the observations, which are known to have been recorded with care. The method of seismic sounding has been widely used in Antarctica to measure the depths of ice.

GRAVITY MEASUREMENTS

During the I.G.Y. the measurement of gravitational force which varies from place to place on the surface of the earth has been extended to places where no observations had hitherto been made, as for example in Antarctica. Relative gravity measurements were made by the Commonwealth Transantarctic Expedition during the traverse of Antarctica, and at Halley Bay. With funds provided by the United States, the Department of Geodesy and Geophysics of the University of Cambridge made a series of pendulum gravity observations at eight stations between Teddington, England, and Johannesburg, South Africa, and a further series at stations between Washington, D.C., and South America. From the many observations it will be possible to improve our knowledge of the relations of the gravity network of the different continents.

It has been shown by R. H. Merson and D. G. King-Hele (*Nature*, 1958, **182**, p. 640) that scientific observations of the artificial earth satellites can be used to investigate the earth's gravitational field. They deduce that the earth's flattening at the poles is less than previously believed, by about one part in 300. They find that variations in the period of revolution of several satellites suggest that there are irregularities in atmospheric density from day to day and over longer periods, with some evidence of a 27-day period, probably of solar origin.

Professor Kozyrev of Pulkovo Observatory has found that th

earth is flatter in the northern than in the southern polar regions, and claims that the rapidly rotating planets Jupiter and Saturn show a similar asymmetry. If this applies also to the sun, it might possibly account for some of the unexplained anomalies in the progress of sunspots across the earth's disc.

NUCLEAR RADIATION

The network of nuclear sampling stations set up during the I.G.Y. gives some clear ideas of the background of such radiation in the atmosphere. Observations made immediately after the Windscale nuclear reactor accident in October, 1957, brought to light some radically new features of atmospheric dispersion.

THE ANTARCTIC

Some features of the investigations in Antarctica have already been mentioned. In all, 57 stations have been established there, by twelve countries. In addition, the transantarctic expedition made many scientific observations, including seismic measurements of ice thickness, on its way across the continent.

The investigation of the structure of Antarctica has, of necessity, been 'spotty', owing to the difficulties of travel. It is still not definitely established whether Antarctica is one continent, two continents or an archipelago covered by ice. This will undoubtedly be readily settled when all the seismic and other investigations of the surface are examined and considered. Over 50 years have elapsed since Dr. Griffith Taylor, who was geologist on Scott's first expedition, put forward the idea that East and West Antarctica were separated under the ice by a transcontinental trough, joining the Ross and Weddell Seas. Eastern Antarctica, which lies mostly to the east of the 0°–180° meridian, is a vast elevated 'shield' of pre-Cambrian rock over 500 million years old, while Western Antarctica is characterized by folded ranges and plateaux similar to those of the Andes, of which they are thought to be an extension, of much later date than Eastern Antarctica.

Seismic data gathered by American ice-traverses point to the same conclusion. The bottom of the trough averages nearly 4 000 ft below sea-level. Further seismic investigation should readily decide the question whether Antarctica is one continent or two, or an island archipelago.

In a letter published in *Nature* (31st January, 1959), three New Zealand scientists, F. F. Evison, C. E. Ingham and R. H. Orr, showed that it was highly probable that Antarctica is a single continent. This was deduced from the crustal thickness, for which, from seismic observations, they gave the value 35 km. Continents generally have crustal thickness of 30–40 km, the corresponding value for coastal regions being 5–10 km.

The surface of Antarctica presents many oddities, among which the so-called 'oases' are noteworthy. The Bunger oasis, at which the Soviet Union set up their third base in late 1956, consists of over 300 sq miles of bare rock, interspersed with lakes. No satisfactory explanation has been found for the fact that, in a region such as this, where there are furious blizzards in winter and temperatures barely rising above freezing-point in summer, there are ice- and snow-free rocks and water in the lakes in summer.

THE NEXT STAGE OF THE I.G.Y.

The I.G.Y. came to an end on the 31st December, 1958. The exchange of information between the nations has been highly successful, and a vast accumulation of observations has been gathered into the World Data Centres. The next stage of our task will be the reduction and discussion of this accumulation of information. We in this country have taken firmly the view that we shall now give first priority to this task of reduction, interpretation and correlation of the observations we have accumulated.

FUTURE INTERNATIONAL CO-OPERATION

A meeting of I.G.Y. scientists was held in Moscow in the summer of 1958. The representatives of all the nations at this meeting were unanimous in regarding the I.G.Y. as having proved to be outstandingly successsful, so much so that it was desirable to continue the observational and data-collecting work for at least another year. This continuation of international activity will not bear the name 'I.G.Y.', but will be known as 'International Geophysical Co-operation 1959'. During this additional year of co-operation, the prescribed schedules of the I.G.Y. will be abandoned, and each nation will be free to discontinue any of the special investigations made during the I.G.Y. and to choose for itself those studies which it wishes to continue, and the level of concentration on those studies, as well as being free to initiate any studies which it may wish to take up.

Many of the routine observations which were intensified during the I.G.Y. form part of regular scientific services, such as in this country the Meteorological Office, the D.S.I.R. Radio Research Organization and the National Institute of Oceanography. These observations are likely to be continued, though not of necessity at the same level of activity as during the I.G.Y.

The Special Committee for the I.G.Y. (C.S.A.G.I.), which was set up by the International Council of Scientific Unions to plan and co-ordinate the I.G.Y., will cease to function on the 30th June, 1959. A new special Committee for International Co-operation in the Geophysical Sciences (S.C.G.) has been established to organize the continuation of the co-operation, so successfully demonstrated during the I.G.Y.

As for the I.G.Y. itself, there remains the Herculean task of reduction and discussion of the observations, and of publishing the results, so that they may be available to all who desire to learn of them. It must not be assumed too lightly that the observations will in all cases find easy explanation: it is probable that many will be correctly explained only when several generations of geophysicists have considered them.

The greatest contributions to the I.G.Y. have been those of the United States and the Soviet Union, especially in the exploration of conditions in the upper air, with instruments carried by rockets and satellites.

The U.S. contribution in this and other fields has been summarized in papers written by Dr. Hugh Odishaw, published in the American journal *Science* during 1958 and 1959, and I wish to record my debt to these papers.

681.14

The Institution of Electrical Enginee
Report No. 315
Sept. 195
©

MEASUREMENT AND CONTROL SECTION

SPECIALIST DISCUSSION MEETINGS ON NEW DIGITAL-COMPUTER TECHNIQUE 16TH–17TH FEBRUARY, 1959

SESSION 1.—CHARACTER RECOGNITION

INTRODUCTORY LECTURE

By C. E. G. Bailey, M.A., Member.

The character-recognizing device, or 'reader', is basically a data-processing machine. In the course of the history of its invention, various forms of output have been proposed, including audible or tactile signals to enable the blind to read and the automatic punching of telegraph tape. At present, interest is concentrated on feeding a computer, either directly or via some form of intermediate storage. This interest has been greatly stimulated by the need for commercial data-processing, and development is proceeding widely and intensively.

The human eye and mind recognize characters by the general disposition of parts relative to the whole, and the geometric positioning of white and black in a character is of little importance. This form of 'morphological' recognition is difficult to parallel by means of a machine. On the other hand, the features of handwriting, and to a lesser extent those of print, are organized to a morphological rather than a geometric specification. Early inventors of readers ignored all but the geometrical approach.

The Optophone, patented in 1914 by Fournier d'Albe, produced distinguishable audible tones from the white areas of each of a number of horizontal scanning lines. It was designed to enable the blind to read. The more useful complementary system, in which the presence of a tone indicated a black area, was patented in 1919. A device, based on the known method of automatically punching cards for Jacquard looms, which gave tactile representations of the letter being scanned, was proposed in 1926.

The first readers aiming to produce a single output signal for each character utilized a template method. The black outline was optically superimposed on a number of standard forms in succession, a signal being given to indicate congruence. Difficulties of registration, common to most geometric systems, arise with this method, but it might well be revived using, for example, image dissectors and rapidly whirling the electronically-reproduced image over the template.

None of the early inventors paid attention to the problems of mutilation arising when typewritten matter is scanned. They likewise glossed over the complexities of conventional type-faces. But two new principles were established in the 'thirties and 'forties which have contemporary importance, namely the restriction of scanning to selected areas and the development of special type-styles designed to be easily recognized by automatic means. These principles were incorporated in some devices by first selecting test areas with a stationary mask and then scanning them with a rotating mask. In later developments, special type-founts were devised to enable all the test areas to lie on a straigl line. In another device the fixed and moving masks wer replaced by a single mask carrying out a special movemer over a succession of stations.

In the 'forties and 'fifties certain techniques in other fielc directed the minds of inventors. The television raster scan ha been known for some time, but the arrival of pulse-length di criminators and storage devices enabled them to scan tl character as a whole and then defer its dissection and recognitic until it had been dissected into pulse lengths and the resul stored.

Zworykin devised a method of horizontal scan using puls length discrimination. An American corporation use vertic scanning and a wide variety of techniques to scan stylize numerals; their design, first among those mentioned, has record of useful commercial application. More recently, a English company has demonstrated a device—'FRED'—suitab either for optical or magnetic scanning of stylized numerals.

Recent American developments for differentiating betwe the names of the States of the Union aim at covering a wi variety of type-founts and parallel some features of *Gest* recognition by their use of combinations of letters.

The problem of reading handwriting has baffled most invento A nearly morphological approach has now been patented which a horizontal sawtooth scan is interrupted when it strik the left- and the right-hand boundaries of the character. T positions of these intercepts may be regarded as a contour wi respect to the vertical component of the scan. Sharp disco tinuities occur with various figures; the presence and sign these discontinuities constitute a code for the figure, which evaluated by means of a matrix store. To be readable aut matically the handwritten figure must fulfil a certain specificatic which imposes some limits—not very severe—on the writer.

Even fewer limits are imposed by the figure specificati developed by Dr. Dimond, in which the intersection of t figure with any of seven construction lines is observed. Dimo developed his specification for use with conductive ink, but early application by Naumberg points the way to performi this operation optically. Apart from this proposal, the ma development effort had been concentrated on optical scannin

Within the past few years, however, a great deal of work l been done in America on the recognition of characters prin in magnetically-permeable ink. The basis of this method is read the effective area of magnetic ink under a horizonta narrow head-gap covering the figure from top to bottom.

Mr. Bailey is with Solartron Electronic Business Machines, Ltd.

The present state of the art can be illustrated by reference to the ERA, which is now in the production stage. This machine is designed to read typewriting or ganged printing with low throw-out and error rates. This material differs greatly from the ideal clean-cut black-on-white typography with which the early inventions described were concerned. The major effort has been spent in solving such secondary problems as mutilation and misalignment at least as much as in basic character recognition. In order to achieve speed and flexibility of movement from one character or block of characters to another, flying-spot scanning is used. Each character is covered by a small raster, and many such rasters can be accommodated in the tube face. The scanning waveforms thus cover a wide band of frequencies, including zero, and consequently electrostatic scanning was chosen. A modified cathode-ray tube has a spot size small enough to accommodate readily 15 adjacent characters in the dimension at right angles to the paper movement. There is little to spare in signal/noise ratio: two photo-multipliers are used for pick-up with white-clamped automatic gain control.

The angle of scatter which gives optimum signal/noise ratio is not critical. With unglazed paper this is independent of angle of incidence from 0° to 60°, but with glazed paper and certain inks there is too little contrast from 0° to 30°, so a scatter angle of about 45° is used.

Each character is submitted in general to three scans or 'frames' with vertical lines. The first of these records the density of the character, and operates a clamped control to ensure that the two following frames have favourable black/white contrast, differentiating the true character from the smudges or 'halo' which surround it. The character is then provided with a fairly clean black edge for the second frame, which covers the same area and measures the extreme black edge at either side, top and bottom of the character. A clamped control derived from this is applied to centralize the character, which it does to within $\pm\frac{1}{2}$ element vertically and horizontally during the third frame. The third frame, with nearly correct contrast and registration, is used for reading.

Time is economized by coarsening the line pitch during the contrast-measuring frame, and by narrowing the actual width of the third or reading frame. The latter consists of 12 lines, each of which is divided into 14 vertical segments, thus making 168 elements or cells in all. The light value of each cell is integrated and a black or white decision is stored in a cell core. (This core is duplicated for black and white as a parity check.)

The timing circuits are performed by transistor-driven 'and' gates from clock pulses and counters. Much care has gone into using the high-frequency transistors in low-power driving positions. By interleaving the clock pulses, the writing speed has been doubled for a given reading speed.

The store is connected for analysis to the 'logic'. Organization of the logic is complex and can be indicated only generally here. The character, which after the registering process still may appear in any of four positions, may be divided into 'features'—those components useful to distinguish one character from another. The lowest element in the hierarchy of the field is a cell; cells are organized into 'domains', typically of three cells, which are joined in 'and' or 'or' connection. In general, 'or' connections are used for black domains, and 'and' for white. Domains are connected, either singly or in blocks, to make a 'statement' about a character in one of its positions. These statements, in Boolean terms, may be as high as the sixth degree. In general, statements about a character in one of the four permitted positions are 'or'-gated to the final character statement.

This organization results in a very low error-ratio; the main difficulty with some documents is likely to be the throw-out ratio. Investigations are being made into the possibility of replacing the permanent wiring in the logic by a coded set of signals on, say, a magnetic drum store, which should result in the appropriate matching patterns being offered serially or simultaneously to the filled store.

The first applications are to continuous-roll reading. To minimize the on-line loading time of the rolls, a feed mechanism has been devised taking pre-loaded cassettes. When a cassette is inserted, rollers, followed by a vacuum feed, pull the paper past the scanning position, and the roll is automatically spooled up.

The spacing between lines of print varies in practice between one originating machine and another. The smoothed output of the vertical misregistration signal is accordingly taken to a servo system which accelerates or retards the paper feed accordingly.

A SYSTEM FOR THE AUTOMATIC RECOGNITION OF PATTERNS

By R. L. Grimsdale, M.Sc., Ph.D., Graduate, F. H. Sumner, B.Sc., Ph.D., C. J. Tunis, B.Eng., M.Sc., Ph.D., and T. Kilburn, M.A., Ph.D., D.Sc., Member.

There is an important need for machines which can read. This arises in many fields, particularly in commerce, where such operations as the sorting of cheques and mail are now being mechanized. In order to make these schemes fully automatic, machines are required which can recognize symbols written on paper. Further applications are suggested in the reading of documents in preparation for language translation by a digital computer. The idea of reading machines is not new, and had its beginnings in attempts to make devices to help blind people to read books. One such device produced an audible signal as it was moved across a line of print. This device was, however, more in the nature of a transducer rather than a recognition device, because the recognition had still to be done by the user.

The authors are at the Electrical Engineering Laboratories, Manchester University.

A range of recognition machines have been devised or constructed and the majority of these employ a comparison mechanism. One scheme uses a rotating disc with a number of apertures cut to the shape of the letters of the alphabet, numerals, etc. As the disc turns, an arrangement of lenses and a lamp cause images of the letters on the disc to fall upon the unknown character. The amount of light reflected is measured by a photocell, and recognition occurs when the amount of light is a minimum. Such a system has a number of obvious disadvantages. The figures to be identified must closely resemble those on the disc, and they must be correctly positioned and have the correct angular orientation. A number of electronic systems have been devised which, while being improvements over the disc system, still have its major disadvantages.

In contrast, the present system is for the recognition of

'patterns'. Character-recognition devices serve only to distinguish between a limited set of characters, but patterns are recognized as an assembly of lines and curves. The system obtains a description of the number of lines and curves in the unknown pattern, together with details of how the parts are joined. Such descriptions are invariant to changes in position, vertical, horizontal and angular, to change in size and to small changes in shape. To demonstrate the system the whole mechanism has been simulated on a medium-speed computer, although work is now in progress on the development of special-purpose equipment to perform certain of the operations at greater speed.

The pattern is presented to a flying-spot scanner which transfers it to the high-speed store of the computer, where it is represented on a raster of 40 × 64 points. The pattern is examined by a programmed scan which results in its being divided into a number of parts, or 'groups'. These groups may not necessarily coincide with the most convenient or natural divisions of the pattern, particularly if it is a letter of the alphabet presented at an angle. Special provisions are made to allow for imperfections in the figure such as irregular edges, breaks, or dirt on the paper. The groups are described in terms of their length, position, slope or curvature, and are then assemble into the natural pattern divisions.

The complete pattern is described by a 'statement', whic gives a list of these natural divisions and the way in which the are joined. The statement, which is coded in the form of a ro of binary digits, is effectively a one-dimensional pattern corre sponding to the two-dimensional pattern presented to th flying-spot scanner. The statement is compared with a set o statements stored within the machine; this comparison make use of a scoring system, and the name of the pattern whos statement gives the greatest score is printed out. When ther is confusion, this is indicated by the machine and the relativ probabilities of the possible patterns are printed out.

If the machine is presented with a new pattern, it will indicat that it does not know that pattern, and the name of the patter may then be supplied to the machine by the operator. O subsequent occasions the machine will recognize this pattern even though it is drawn afresh by hand. This ability to lear is an important feature of the system. The machine build up its list of standard patterns in this manner and is thu universal; it is in no way restricted to any particular type face o style.

THE AUTOMATIC RECOGNITION OF TYPEWRITTEN NUMBERS

By W. Dietrich, D.Ing.

While the magnetic character-readers currently available are insensitive to paper fouling, they do not supply all the information a character contains. In contrast, an optical system reproducing the shape of a character electrically supplies for the identification practically all the information the character bears. Hence this system can recognize a large number of characters of various shapes, e.g. both numbers and letters. On the other hand, special precautions are necessary to balance the effect of fouling, i.e. the smudging of paper or characters. This contribution deals with such an optical method, with special emphasis on the centering of characters and the identification of fouled characters.

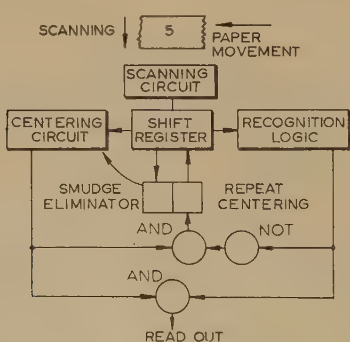


Fig. 1.—Block schematic.

A block circuit is shown in Fig. 1. The characters are moved in a direction perpendicular to the scanning sweep, so that the character is gradually dissolved in an area. Under the condition that the character becomes stored in the shift register synchronously with the scanning process, the registered information will correspond to the quantized but natural shape of the character. When the registering process is stopped, the register stages wil retain their information, which then can be made visible, a shown in Fig. 2, with the aid of small lamps connected to th stages.

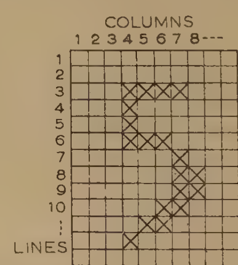


Fig. 2.—Shift register and lamp field.

Now the centering circuit must centre the information in defined area of the shift register, say, in the top left-hand corne To do so, the whole information is shifted by pulses, first to th left and then to the top until some stages of column 1 and line report 'black'. For this shifting procedure the stages of th register are connected in a 2-dimensional manner.* The centere character will generate a centering signal. If the logic identifie the character at the same moment, this character is read ou through the output circuit via the 'and' gate (see Fig. 1).

The following examples are limited to arabic numerals a produced by a typewriter. However, these numerals are in n way changed in their original shape. A clean numeral, suc as the 5 in Fig. 2, can be easily centered and recognized. Th real task of the character reader, however, is to recogniz smudged numerals. Smudge means here, in a wider sense o this word, the variations of a character in consequence of o or new ribbon on the typewriter, smooth or temperament

Dr. Dietrich is with Standard Elektrik Lorenz A. G., Informatikwerk, Stuttgart.

* Steinbuch, K.: 'Automatische Zeichenerkennung', *Nachrichtentechnisc Zeitschrift*, 1958, 11, pp. 210 and 237.

touch, original or first copy, dirty keys on the typewriter and finally, all disturbances occurring from the paper surface.

For the reading of smudged numerals it will be more suitable, not only to use the first column and the first line as centering criteria as mentioned above, but to supervise electrically the whole area where the centered character shall be located, e.g. 6 columns and 10 lines. Thus we may say: Any information will be signalled as centred only if it is of the approximate size of a numeral. This centering method will eliminate many smudges as well as other unwanted printing, such as the printed horizontal and vertical lines which designate a writing space on a commercial form.

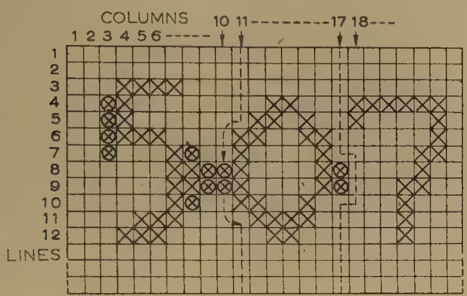


Fig. 3.—Number 507 affected by smudge.

☐ White.
☒ Black.
⊗ Smudge points.

Fig. 3 shows the number 507 fouled by some spots of smudge. To simplify the discussion only a few spots are assumed, but these have a particularly adverse effect. It may be mentioned that the shift register must not be wider than a character. The wide register of Fig. 3 serves only to show more numerals and how they may be connected.

First, the numeral 5 in Fig. 3 will not be properly centered because of the smudge points in the third column. The solution of this problem is based on the fact that the numeral will not be recognized by the logic. Information which is centered but not recognized generates a 'repeat centering' pulse (Fig. 1). As a result, the information is shifted by one column to the left and recentered. This action may also be done in a vertical direction.

The centering conditions and the repeated centering method are applied with the aim of eliminating smudge and of centering smudged numerals as if they were clean. The advantage of these methods is that the design of the logic then can be based on the ideal numerals, and the rather voluminous wiring and many components used can thus be brought into a systematic order.

Fig. 3 also shows smudge points between two numerals. These are particularly awkward, because the character reader now has the difficult task of locating the beginning of the next numeral. To solve this problem, two different methods are used. First, the quantity of black information in two adjacent defined columns is measured, as in columns 10 and 11 of Fig. 3. If there is more black information in column 11 than in column 10, this is evaluated as a signal for the beginning of a new numeral. This signal is now stored and used as a centering signal after the numeral 0 has been shifted towards the left in the normal course to the first column. Secondly, two defined columns, such as 17 and 18 in Fig. 3, are supervised. If, along the whole length of these columns, any one of two adjacent stages is white, this is again taken for the beginning of a new numeral. Such a way is indicated by the interrupted lines in columns 17 and 18. Both of these methods can be modified if needed. Since they operate in a parallel manner, it will suffice if only one of them will discover the beginning of a numeral. Furthermore, it will be sufficient to find this beginning only approximately, since the 'repeat centering' circuit operates additionally and may centre the information correctly.

A device for reading ideal characters needs additional developmental efforts if it must deal with the difficulties occurring in a practical application. This contribution shows how an essential problem, i.e. the smudge (in the widest sense of this word), may be overcome to some good extent. When using the shift register as a store for the scanned character, further methods may be accomplished in a similar manner. Other methods may be adopted with the aid of the logic circuits and an analogue grey control. All these methods will contribute to the overall advantage of an optical reader, which permits the storage of a character in its natural shape, thus supplying for identification practically all the information the character bears.

PATTERN RECOGNITION BY MEANS OF AUTOMATIC ANALOGUE APPARATUS

By W. K. Taylor, Ph.D., M.Sc., Associate Member.

For the purpose of this discussion the word pattern will be used to denote any n-dimensional intensity distribution which can be measured by means of a transducer system. Characters are usually relatively simple two-dimensional intensity distributions and can be measured to any desired degree of accuracy by a mosaic of photosensitive transducers. The output signal of each transducer can be made proportional to the average light intensity over a small element of the pattern. Recognition of patterns is equivalent to classification of sets of transducer output signals. When the patterns are fixed in position relative to the transducers, so that maximum or zero signals appear at each transducer, the classification is easily achieved by a network of mechanical or electronic switches, since the signals are in a binary code. A well-known example of this special binary pattern is the 5-hole punched tape.

Binary patterns are classified correctly only if all the digits are read without error by the transducer. If one digit is in error, the remaining correct digits do not have any power to produce a correct classification. This is not true when characters made up of a large number of dots are recognized by a human operator. If, for example, a capital letter T is made up of nine dots, five for the horizontal and five for the vertical stroke, it is still possible for the operator to classify it correctly if several dots are removed, provided that they are not all removed from the same region of the character. In other words, it can be said that the remaining dots still have some control over the classification. The simplest assumption which can be made to account for this phenomenon is that all dots have an equally effective influence on the classification, irrespective of the actual number present. This assumption requires that the principle of superposition should apply and that the signals should be added in a linear analogue network rather than made

Dr. Taylor is in the Department of Electrical Engineering, University College, London.

to open and close the gates of a switching network. By operating with the sums of sets of signals, we obtain the additional advantage of not having to select an arbitrary clipping level above which a signal becomes '1' in the binary digital system. This property is particularly important in the general case when patterns do not occupy predetermined positions on the transducer mosaic, since the edge of a shape may then cover any fraction of the light-sensitive area of a transducer.

There are 2^N-1 ways of selecting sets of N signals to form sums of sets, and a method of distinguishing between the sums is required. The largest sum is produced by the largest pattern, but if each sum is weighted according to the number of signals it contains, it is possible to arrange for every possible pattern to produce the largest weighted sum at a particular adding unit, irrespective of the overall pattern intensity level. This means that a dimly illuminated character will automatically be given the same classification as the same character in brilliant illumination. This again appears to be a human-operator characteristic.

When the number of transducers required to sample the intensity distribution becomes large, it is necessary to find ways of reducing the number of sets of signals to be handled. This can be done by eliminating signals which correspond to areas containing no detail and by restricting the number of signals in the sets. The remaining signals must be sufficient to control a classification system with the required accuracy. In principle, it is possible to connect the adding units to a set of output units so that patterns are classified according to the designer's specifications. There may, however, be applications in which the classifications are specified by an unskilled operator or by a second machine. It is then necessary for the recognition apparatus to be capable of 'learning', or of storing the classifications in a memory system. This has been achieved by making provision for all possible classifications and giving the various signal transmission paths zero gain initially. The classification signals then increase the gains of the paths taken by the pattern signals until the patterns are correctly classified. After this automatic setting-up procedure has been carried out, the path gains can be measured and the variable-gain units replaced by fixed attenuators to produce a fixed-purpose pattern-recognition machine. The operating speed of the machine is limited only by stray capacitance, and by adding unit resistors sufficiently small, reading speeds of 10^6 characters/sec can be achieved.

RECOGNITION OF NUMERALS BY CONTOUR FOLLOWING

By W. Sprick, Dr.rer.nat., and K. Ganzhorn, Dr.rer.nat.

Arabic numerals present the main amount of recognition information at both sides as seen horizontally from right and left. The movement of a pencil writing a numeral changes its direction rapidly at the points of reversal, breaking, starting and ending, which may be considered as the essential characteristics of every numeral.

By means of a contour follower it is possible to find voltage functions analogous to both sides, e.g. if the numeral is scanned from top to bottom. The characteristic points are then detected by differentiating these functions and transforming them into a characteristic differential pulse pattern. The comparison between the received pulse pattern and the reference pattern must fulfil the fundamental conditions that no time coincidence and no amplitude equality are used as criteria. These conditions, in association with the differential pulses, warrant a certain invariancy with regard to the position, size, inclination and some distortions of the figure, which means that it is possible to cover a wide variety of type styles by a single reference pattern without the need of positioning or centering.

The contour follower is simulated by a circuit arrangement which scans the numeral in horizontal lines from the top to the bottom, as in the common television system. It converts the pattern of the video pulses to two patterns of pulses, the envelopes of which are equal to the right- and left-hand sides of the numeral. The conversion from video pulses to both patterns of envelope pulses is obtained by clipping the video pulses at the momentary value of two sawtooth voltages, one being in synchronism and the other reversed with regard to the scanning voltage.

The differential quotient is obtained by means of an *RC* circuit, with a small time-constant, which delivers sufficiently evaluable differential patterns.

Dr. Sprick and Dr. Ganzhorn are with I.B.M. Deutschland G.m.b.H.

The comparison—or, more generally, the evaluation—circuit comprises an arrangement of criteria detectors, storage units which store the criteria during the scanning procedure and a co-ordination logic. This logic performs the connection between the criteria and the output lines associated with the numerals 0–9 by means of the reference-pulse patterns, realized by 'yes' and 'no' connections in a co-ordination matrix.

The first test model was designed to demonstrate the versatility of the system. It does not seem an essential restriction to divide the scanning area for a numeral in three horizontal zones. The following features were used in combination as criteria: the polarity of the differential pulses (+ or —), the zone (1 or 2 or 3) and the function of the left- or right-hand side. It is demonstrated that the number of types styles which is covered by a single reference pattern is surprisingly high.

Subsequent efforts were directed to the elimination of the zones. Although it seemed to be possible to make the zones dependent on the figures themselves, this was not pursued. This solution was achieved by increasing the number of the criteria and their modes so far as possible with regard to the fundamental conditions not to use the time coincidence and the equality of amplitude as criteria.

It was found particularly successful for scanning printed numerals to introduce, in addition to the differential pulses, the checking of long or short strokes in the horizontal direction and the counting of the number of crossings in one horizontal line.

Details are given in the following patents:

German patent No. 953474, November, 1956.
German patent No. 961225, March, 1957.
U.S. patent No. 2738499, March, 1956.
U.S. patent No. 2838602, June, 1958.

DISCUSSION ON CHARACTER RECOGNITION

The discussion was opened by **Mr. R. H. Tizard**, who suggested that instead of classifying character recognition systems into those of a morphological and those of a geometrical approach, it might be reasonable to classify them in terms of the amount of information which was rejected before the recognition process began. It was pointed out that, if we are to come anywhere near to matching the ability of the human eye and brain, we shall have to make use of all the information available, and it was suggested that the future might well lie very much in the use of probability techniques in which the probability parameters were determined by a learning process. Mr. Tizard supported his remarks with an account of some programming he had been carrying out in which he had started with reading up to eight characters, but he did not think that the complexities would increase very rapidly with an increase in the number of characters to be read. The first development was what he called a negative process. The machine was told, for example, that 'This is E not F'. The next stage was a self-classifying process in which the machine, given certain characters, would improve the basis of this distinguishing to get the best possible results. The programme employed a method in which scoring was used and in which there was a criterion of the excellence of the results.

Mr. Tizard also spoke about the possible automatic reading of handwriting, and said it was certainly the case that the reading of ordinary manuscript was impossible on a letter-to-letter basis, and that the human being succeeded by word recognition and a great deal of use of context. Until a machine capable of doing this sort of thing was produced, it was very much star-gazing to talk about the automatic reading of manuscript.

Mr. C. Strachey, who spoke next, had a word to say about the use of the word 'learning'. He suggested that there were two distinct processes involved in the machines described which, in the course of time, gave an improved performance. One process was for the machine to alter its basis of selection, and it was possible to arrange for the machine to do this automatically. However, he considered this to be not a process of learning, but a process of adjustment.

Mr. N. A. F. Williams spoke about the speed with which character recognition machines were likely to operate and asked for further information. **Mr. J. C. Selman** and **Mr. E. A. Newman** also spoke about the size of the equipment and costs, and the latter suggested that it would be interesting if the costs could be expressed in pence per correct character read per second.

Mr. R. Hayes and **Mr. D. M. Taub** also spoke about costs, and in particular the cost of the character-recognition machine in relation to a digital computer. Mr. Taub inquired whether it was cheaper to recognize a character which was originally designed to be recognized by the human eye than to incur the expense of more complicated printing machines or typewriters which would print in code, which the machine could then recognize more easily.

Mr. C. E. G. Bailey thought that classification as proposed by Mr. Tizard by the amount of information thrown away was valuable, and he went on to speak about the design problem of not throwing redundant information in the wrong order and so getting a poor signal/noise ratio. He next answered Mr. Williams by stating that in the pilot model of his machine the speed was 120 characters/sec and in the present model it was 280 characters/sec. If his mental arithmetic was correct, the cost of the machine was 24 000d. per character read, which was increased in the ratio of 1·0001 : 1 if he had to give the cost per correct character read.

Dr. R. L. Grimsdale followed Mr. Bailey and mentioned that in their system a digital computer was used and therefore the cost question could be answered by saying that, if one had a digital computer available for other purposes, the cost of their character-recognition system was not large, but if a separate digital computer had to be purchased, it could be large. He agreed with Mr. Strachey regarding the use of the word 'learning', and in answer to the questions about operating speeds he mentioned their hopes of obtaining speeds of 100 characters/sec.

Dr. W. Dietrich gave some information regarding the size of his equipment and mentioned that it recognized figures at a speed of 50 figures/sec.

Dr. W. K. Taylor agreed with Mr. Tizard's remarks that a digital computer could be useful for making statistical analyses of handwriting or similar characters, and then spoke about postal addresses being written in a code form. He also spoke about the speed of working of his system and explained that extremely high speeds of working were theoretically possible, the limit being obviously set by the input equipment.

Dr. K. Ganzhorn wound up the discussion by saying a few more words about his contour-following principle, and mentioned that his apparatus was operating at speeds of the order of 100 characters/sec. It would seem then that all the different methods described were capable of operating (or could operate) at substantially the same speed.

SESSION 2.—PERIPHERAL EQUIPMENT—I

INTRODUCTORY LECTURE

By B. W. Pollard, M.A., Associate Member.

The main types of equipment available for peripheral use with digital computers will be described and the general requirements summarized. Consideration will be given to the performance of existing equipment and to the likely requirements in the future. There will be a discussion of the problems facing the user of digital computers and his need for a simplification of the methods of presentation of information to a computing system.

There will be a discussion of the methods of attachment c peripheral equipment to computers, together with suggestion for simplifying such methods. Proposals will be made for th standardization of some of the characteristics of items c peripheral equipment to reduce the duplication of developmen effort and increase the availability of equipment. In particula the need for intensified effort in the development of items c peripheral equipment will be stressed.

THE DESIGN OF HIGH-SPEED PHOTO-ELECTRIC TAPE READERS

By M. V. Wilkes, M.A., Ph.D., F.R.S., Associate Member, and D. J. Wheeler, M.A., Ph.D.

A paper-tape reader intended for general use in the input channel of a digital computer should allow rows of holes to be read individually or in rapid sequence as called for by the programme. High reading speeds are required, and, at the same time, the facility is desirable of stopping a tape moving at high speed and restarting it without losing any of the punche information.

The contribution will consist of a general description of th problems encountered in the design of such a tape reader and description of a machine capable of reading tape at the rate c 1 000 holes/sec which has been developed in the Mathematica Laboratory at Cambridge.

Dr. Wilkes is Director of, and Dr. Wheeler is in, the University Mathematical Laboratory, Cambridge.

TAPE DRIVING AND SPOOLING EQUIPMENTS

By B. G. Welby.

A decade ago, punched-paper-tape readers were primarily designed for application in teleprinter systems and consequently operated at speeds of less than 10 characters/sec. The development of electronic computers produced the need for faster reading speeds coupled with a reading system which would be both reliable and non-damaging to the tape, even after repeated passage through the reader—the latter requirement also dictated that an alternative to the sprocket drive for tape feeding should be used. One of the earliest readers specially designed for use with electronic computers has a speed of 200 characters/sec and uses a photo-electric system of reading, the tape being driven by friction from a tape-feed mechanism which makes use of a differential gear assembly in conjunction with a pair of brakes.

As the design and application of computers progressed, larger quantities of data at higher speeds of reading were required. Initially this need was met by providing separate spooling devices which could be used in conjunction with small table-top readers. This arrangement is very wasteful of desk area and is unduly complicated for the threading of tape. The present state of tape-reader development is to combine the spooling with the reader into a single cabinet, which, of necessity, resembles a magnetic-tape recorder of similar speed and capacity.

A punched-paper-tape reader has recently been developed for speeds up to 1 000 characters/sec and incorporates spooling facilities. In the design of this equipment several novel features have been developed, and these will be described. The design objectives have been to produce a reader which will accept any of the three standard widths of tape up to eight channels wide and be capable of stopping the tape within a character displacement, even from the full speed of 100 in/sec. Because tape lengths may vary from a few inches to over 1 000 ft, provision has been made for inserting short lengths of tape without the use of spools, while longer tapes contained on spools can b loaded in a simple manner.

Unspooling and respooling is accomplished by servo-drive spools situated on either side of the reading head. A vacuu tape-reservoir box interposed between each spool and the hea provides tape storage to meet the immediate demands of th high-speed tape drive. This enables the spool to accelerate mo slowly than the low-inertia capstan, which must accelerate rapidly as possible. The amount of tape contained in eac reservoir box is monitored by photo-electric cells in conjunctic with a simple illumination system designed to produce a line response to the quantity of tape stored. After amplification, tl cell outputs energize field windings on the respective servo moto

Signals from the reservoir photocells are amplified by a d.c connected transistor amplifier which incorporates an error-ra circuit to give the required servo damping. To minimize pow dissipation in the final amplifier stage, a circuit arrangeme whereby the final transistor is either fully conducting with collector bottomed or completely cut off is used. This is achiev by incorporating two feedback loops, one positive and the oth negative. The positive feedback is an interconnection betwe the last two stages of the amplifier to ensure that the outp transistor is either fully on or off. The negative feedback derived from a low resistance in the load circuit to produce voltage proportional to the current flowing, and the signal th derived is passed through an *RC* network to obtain phase shi and is then injected into the amplifier input, together with t photocell signal. Owing to the phase shift applied to t negative feedback, the circuit is oscillatory, with the final sta flipping between on and off with a mark/space ratio determin by the photocell input. With an oscillation frequency which high compared with the time-constant of the servo-mo windings, the circuit produces a very good linear amplificati of current with a controlled-power/dissipated-power ratio about 50 : 1.

Mr. Welby is with Ferranti, Ltd.

Tape movement past the reading head is controlled by a tape-feed mechanism which is capable of accelerating or retarding the tape in 1 millisec from a speed of 100 in/sec. To achieve this performance a low-inertia drive unit comprising a motor-driven clutch, brake and tape-driving capstan has been developed. In construction the brake and the clutch are similar, being electromagnetically operated friction brakes in which the output member is a thin steel disc splined onto the capstan shaft. Each disc is normally free to rotate between the pole faces of an annular electromagnet and a corresponding armature. In the unenergized condition a light spring pressure keeps the armature and disc in contact with the magnet pole-faces; energization causes the flux to pass through the disc to the armature, so that in attracting them together a considerable pressure is developed between the mating surfaces. The disc is therefore clamped by friction to the magnet and armature assembly.

Problems in the design and construction of this drive unit will be given.

A HIGH-SPEED PAPER-TAPE PUNCH

By F. L. Turner.

Since the earliest days of electronic computers, punched paper tape has been regarded as a useful medium for the storage of information for processing in the input and output stages of computing and data-processing systems. This has been due to its inherent advantages, representing as it does a compact and continuous medium of low cost, providing a visible and durable record. In addition, since punched paper tape is also the medium used in telegraphic communication, a 'common language' is available for the direct integration of computation, data-processing and communication systems.

It is not surprising also that conventional teleprinter equipment designed to use this medium has, in slightly modified form, operated adequately in the role of tape producing and editing equipment for computers for some time. Their relatively low speed of operation and their use of a serial-mode transmission code, however, have made them generally unsuitable for direct coupling to computer outputs. It was not difficult to modify telegraph punches to operate from a parallel-mode code and thus permit the simplification of their drive circuits and a small increase in the speed. However, the severe restriction on computer operation caused by the low speed of these machines demanded faster punches if the advantages of the use of paper tape as a storage medium were to be exploited to the full. To this end, the last five years have seen the design of a variety of paper-tape punches which have not only increased the speed available from seven to upwards of 50 characters/sec, but have also permitted the punching of 6-, 7- or 8-track tapes.

While these developments improved the overall position, the gap between even the highest of these speeds and the potential speed of the computer was still so great that attention had to be turned to the design of punches capable of very much higher speeds. What speed should be aimed at was difficult to assess, since the only guidance available was that it should be as high as possible. It was therefore decided to design for the maximum speed practicable with known techniques, compatible with the high order of reliability needed for computer operation. This was evaluated by a series of tests to be of the order of 300 characters/sec. Accordingly, a machine has now been produced to operate at this speed, on either a start-stop or continuously-running basis and capable of punching 5-, 6-, 7- or 8-track tapes when provided with a parallel-mode input.

Despite the emphasis which has been placed on reliability as a design requirement, it has been considered desirable to incorporate a check-back facility which reads the perforations in the tape and enables them to be compared with the corresponding input signals. Synchronizing contacts are also provided to enable the punch to indicate to the computer its readiness to receive the next signal. Tape transport through the machine and the character-by-character feed for start-stop operation have received the special design treatment essential to ensure reliability at the very high speed concerned.

Mr. Turner is with Creed and Co., Ltd.

PROBLEMS IN THE DESIGN AND APPLICATION OF HIGH-SPEED CARD READERS

By J. C. Davy.

'High speed' is here defined as a card velocity in excess of 100 in/sec.

The photo-electric sensing of moving cards permits high information output rates in convenient form for computer input without significant card wear or undue noise. However, the information is only presented once, at a low power level, which may necessitate both amplification and storage.

A large number of types of cell are available, but the high information rate and small size required dictate the choice of a germanium photo-diode having a sensitivity $0{\cdot}25\,\mu A/mW/cm^2$ with lamp of colour temperature 2 500° K. This cell has a very adequate frequency limit of 50 kc/s, but has a relatively high and variable dark current and a sensitivity quoted as varying by 3 : 1 from cell to cell. Both dark current and sensitivity are temperature dependent. These latter factors complicate the selection of a suitable load, whether resistive or inductive. The cell mounting block, if of metal, should be firmly connected to a heat sink well away from the light-source mounting, and if it is of insulating material it should be shielded from unnecessary radiation.

Two new possibilities are the silicon cell (photo-voltaic) which is not so temperature dependent but does exhibit variable sensitivity, and the indium-antimonide cell, which is at present rather expensive.

The variations in dark current and sensitivity of the germanium cell make circuit design difficult and call for a light source providing intense and uniform illumination over a length varying from 2 to 8 in with the type of card and feed direction, without undue power consumption. The beam may also be restricted to only 0·014 in in width at card level to give full cut-off between holes. The filament strip-light was tried and found unsatisfactory, since the filament was unduly fragile. A combination of a pre-focus lamp and a parabolic reflector gave a successful source of long life, but with some variation in intensity, owing to the glass bulb and the optics of the parabola.

Mr. Davy is with IBM World Trade Laboratories (Great Britain), Ltd., and was formerly at the Whyteleafe Research Establishment of Powers-Samas Accounting Machines, Ltd.

A further source developed recently uses a long straight wire operated at red heat, with one or two elliptic mirrors, and produces a straight narrow beam of uniform intensity.

The use of a simple commutator has proved impossible at high speed, owing to the existence of a large number of significant and variable factors, e.g. humidity, backlash, manufacturing tolerances, etc. An attractive possibility is to let the card carry its own index marks, but there are significant objections to this, not the least being the resulting incompatibility between existing card files and new equipment. Another approach is the use of contacts opened by the leading edge of the card, but this is very difficult with endwise feed, owing to the close contact pitch required and the drag on the card. A very elegant solution uses clock pulses, generated by the track at a multiple of column frequency and accumulated in a counter whose 'carry' controls a pulse separator. Correct phasing of the 'carry' pulses and the hole signals is achieved by gating the clock pulses with a signal initiated by the leading edge of the card.

The difficulties of designing a high-speed track for moving-card sensing are dependent primarily on the track velocity and the time allowable for presentation of the first column of information to the sensing head after receipt of the 'feed next card' signal. For this latter, a time of 50 millisec is possible but 5 millisec—more directly comparable with magnetic tape—has not yet been achieved. Research into card behaviour at high speed has shown that roll bounce is a significant factor. This can be improved by reducing the mass of the spring-controlled roll. To make successful use of a commutator for indexing—which is economically very desirable—further research will be necessary to improve the performance of both track and magazine.

The silicon cell appears to be the most attractive new development, but many others are likely to become available with current advances in solid-state physics. More efficient light sources will need to be developed. Self-indexing cards and the further development of high-speed tracks should result in a card reader that is very competitive with magnetic tape as a computer input device.

A DESIGN FOR AN AUTOMATIC GRAPH PLOTTER

By M. P. Atkinson, B.Eng., W. T. Bane, B.Sc., Associate Member, and D. L. A. Barber, B.Sc.

Several machines have been designed to plot graphs automatically from digital information; those which incorporate scaling and zero-shift facilities have normally made use of analogue techniques to provide these. The present design uses digital techniques throughout, making the system less liable to accuracy limitations than its analogue counterparts. The increase in complexity of the circuits, however, means that the system is an economical proposition only when use is made of transistors and printed-circuit techniques.

To plot a point automatically, it is necessary to position a plotting head so that its distance from a preset origin, measured along the co-ordinate axes of the table, is proportional to the co-ordinates of the point. The present system, which works in rectangular co-ordinates, achieves this by using a simple digital servo system on each axis, the operating principle of both servo-mechanisms being identical.

The servo principle is based on the use of a counting technique. As the plotting head moves along an axis, electrical pulses are generated at fixed equal-displacement intervals, and are fed to a binary digital accumulator, which is arranged to contain a number representing the distance to be moved. With a scaling factor of unity, each pulse causes '1' to be subtracted from the number in the accumulator. When this number is zero the plotting head has reached the required position and movement along the axis is stopped. At any instant, the number in the accumulator represents the distance to be moved and the sign of the number determines the direction of movement.

No attempt is made to control the speed of movement in accordance with the magnitude of the error indication. The driving motors have three control positions, namely 'forward', 'reverse' and 'stop'. Motion on each axis is stopped by means of an electromagnetic brake, which is operated when the appropriate accumulator registers zero, and a circuit has been designed which causes the brake to develop full torque in 4 millisec.

Scaling factors between $\frac{1}{2}$ and 1, quoted to ten binary places (1 in 1 023), are achieved by subtracting from the accumulator a number equal to the scaling factor for each feedback pulse produced. The pulse is fed in parallel to chosen stages of the accumulator, as determined by the scaling factor. Short delays in the 'carry' lines between successive stages of the accumulator allow this to be done without interference occuring between direct input and 'carry' pulses to any stage. The scaling factor is such that, for any input number, the distance moved is proportional to this number divided by the scaling factor. Factors less than $\frac{1}{2}$ are obtained by multiplying the input number by 2, 4 or 8 during the process of adding the number into the accumulators initially. With this system the maximum possible positional error is just less than two count intervals.

To plot a series of points, the plotting head is first set to the position on the table which represents zero, and the accumulators are set to zero. The co-ordinates of the first point to be plotted are then added into accumulators and the system is allowed to run until both accumulators register zero. The point is then plotted, and its co-ordinates are subtracted from, and those of the next point added to, the accumulators, which therefore contain the difference between the two co-ordinates. The whole cycle is then repeated until all the points have been plotted.

The mechanical designs of the two axes are not identical, although the same operating principles are used. The plotting area is 10 in square. Movement along the x-axis is obtained by moving the plotting table under a fixed gantry which carries the plotting head. In the y-direction the plotting head moves along the gantry, which is set perpendicular to the direction of motion of the table. Both systems are kinematically mounted.

The rotational movement of the driving-motor shafts is converted to linear movement along the axes by means of lead screws. The drive to the table and plotting head is transmitted from the lead-screws by means of shaped wheels which mesh with the screw threads and are attached to the driven member. The wheels are free to rotate about their own axes and also about an axis perpendicular to this, so that the plane of the wheels can take up the angle of the screw threads. Sliding friction between lead-screw and 'nut' is almost completely eliminated by this arrangement, and a small low-torque motor is capable of providing the driving force required.

On each axis the electrical feedback pulses are obtained from two photo-transistors. Light from a fixed lamp falling on the photo-transistors is interrupted by a slotted disc attached to the shaft of the driving motor. This disc has 20 slots equally spaced around its circumference, and the photo-transistors a

The authors are at the National Physical Laboratory.

turned on and off 20 times during each revolution of the shafts. The signals from these photo-transistors are combined in such a way that for each switching cycle a pulse appears on one of two lines depending on the direction of rotation.

The digital accumulators are each made up of 22 identical binary stages. Each stage, which consists of four transistors and associated components connected on a printed-circuit board, can be disconnected easily from its accumulator rack. There are two input lines to each stage, and the line which passes the 'carry' pulse from the previous stage contains a short delay.

Transistors have been used wherever possible in the logical and operational circuits of the sequence control system. It has been necessary to use a few relays as slaves to control operations of motors and brakes, which require high voltages.

With pulses produced at displacement intervals of 0·005 in, the average accuracy of plotting a point is about 0·007 in, although a point can be repeated with greater accuracy. The co-ordinates are presented in decimal form and are converted to binary form during the initial transfer to the accumulators. When the co-ordinates are presented to the machine in parallel, a plotting rate of 3 points/sec has been achieved for points close together.

DISCUSSION ON PERIPHERAL EQUIPMENT—I

The discussion was opened by **Mr. D. F. Walker,** who compared the effect on cost and complexity of plotting tables of the various methods of specifying information to them by means of computers. The first two methods, namely the normal co-ordinate method and the use of first-order difference, appeared comparable. The use of second-order difference would allow a reduction in the cost of producing information from a computer at the expense of a slight increase in complexity of the plotting table, but would seem to offer a reasonable compromise between these two parameters. This system would depend on the use of digital integrators. For systems used in machine-tool control the use of the second-degree incidence method of specifying currents represented a considerable reduction in cost over other methods. Finally, Mr. Walker described a simpler method of converting digital information into linear movement by the use of a motor having a salient-pole rotor and wound armature. Feeding appropriate signals into the armature enabled the rotor to take up various positions.

Mr. M. P. Atkinson replied to Mr. Walker by pointing out that the plotting table designed at the National Physical Laboratory was intended to incorporate the maximum number of facilities. A simpler plotting table incorporating some of Mr. Walker's suggestions might well be possible, although a system based on the use of second-order differences would present difficulties.

Mr. C. Strachey posed three questions, the first regarding the stopping position of perforated tape in the tape reader. **Dr. D. J. Wheeler** replied that the reader stopped within the diameter of one hole.

Mr. Strachey's second question dealt with the arrangements for the control of spooling on the unspooling unit of the high-speed tape punch. **Mr. F. L. Turner** in reply described a variable speed-feed system of the 'potters wheel' type.

Finally, Mr. Strachey asked about the effect of humidity changes on the size of punched cards, and whether this presented any indexing problems. **Mr. J. C. Davy** replied that humidity had a very big effect, but the use of a system in which the card was self-indexing rendered the effect of humidity changes alone insufficient to introduce false indexing.

Mr. W. P. Anderson followed with remarks on the use of computers as parts of large servo systems in process control. Here the digital computer was linked to the machinery to be controlled without intervening human action. He deprecated that the question of the type of link between the servo system and the controlled machinery had not been dealt with in the papers presented.

Mr. Anderson also referred to the possibility of business computers assuming a more decision-making role, with a considerable reduction in the amount of output printing and a resulting requirement for a very small amount of output equipment.

Mr. H. J. Corps took advantage of Mr. Anderson's remarks on servo systems to describe a method of linking digital differential analysers with conventional analogue-computing equipment. The first step was to deal with the problem of converting analogue voltages into a form suitable for feeding into an incremental computer. This could be achieved by feeding the convertor from the analogue voltage, and comparing the rise of current fed through binary-scale transistors at the point of an amplifier also connected to an analogue input current. If the two were unequal, an error signal resulted at the gap of the amplifier driving the transistor in the circulatory group of the binary number. A build-up of this resulted at the rate of one quantum per word until the binary numbers were in relation to the analogue input voltage.

Mr. A. D. Ridlington queried Mr. Davy's claim that self-indexing cards and the further development of high-speed tracks would result in a card reader very competitive with magnetic tape as a computer input device. Mr. Ridlington agreed that high card-reading speeds might be possible, but pointed out that the attendant problems of feeding cards into the reader at an equivalent speed would be formidable. He felt that reading speeds alone did not recommend one input medium against another, and whereas punched cards would remain useful for small or medium amounts of information, he doubted their adequacy as an alternative to magnetic tape for large-scale data-handling operations. **Mr. J. C. Davy,** replying, indicated other factors which should be taken into account when comparing punched cards and magnetic tape, not the least of these being the existence of punched-card files, where transcription on to magnetic tape had little to commend it as opposed to feeding the cards directly into the computer.

Mr. N. F. Fossey requested further information on the nature of the checking device used on the high-speed tape punch. **Mr. F. L. Turner** described this as a mechanical sensing device using peckers which read the holes in the tape when the movement of the die placed the tape against them. The merit of a mechanical as opposed to a photo-electric sensing was the short distance—three characters—between the perforating and reading positions.

In response to Mr. Fossey's request concerning the tape dispenser associated with the high-speed paper-tape reader, **Dr. D. J. Wheeler** indicated first that no servo mechanisms were employed to spool tape directly after it had passed through the reader. A simple basket was used to collect the tapes. For feeding tape into the reader, a simple box of suitable dimensions sufficed, provided that the tape roll did not exceed 2 in in diameter.

For longer tapes the tape was fed over a slightly spring-loaded roller. This prevented tape breakage due to sudden acceleration.

Mr. Fossey's final question dealt with the possible directions of feeding punched cards. **Mr. J. C. Davy** confirmed that, in fact, there were four possible directions of feed, although effectively these amounted to two only. In the broadside direction there appeared to be no virtue between vertically upwards and vertically downwards feeding. For indexing from the card, however, since the cards were printed in reference to the upper edge, this appeared to determine the logical direction in which to feed for sensing. For endwise feeding the logical direction was to the right, so that the reference edge was the left-hand one. This threw on to the indexing equipment the card manufacturing tolerance. In general, however, the advantages of logical presentation outweighed the disadvantages arising from increased inaccuracy.

Mr. M. L. N. Forrest asked about the possibility of using germanium photo-transistors, at least for serial card readers. **Mr. J. C. Davy** agreed with the attractive nature of these devices, but referred to the difficulties of mounting, involving optical spreading methods. Mr. Forrest also remarked on the effect of indexing by selection of one of a number of particular clock tracks, resulting in committal to whichever track was selected for the rest of the 'carry', with the consequent inability to tolerate slip during further progress down the track while it was being read.

Mr. J. C. Davy replied by amplifying his description of the indexing system, pointing out that the clock pulses came from a clock disc and were driven from the track drive. There was no question of a number of particular clock tracks. Nevertheless, there could still be problems of indexing due to slip, provided that this reached sufficient proportions.

Finally, **Mr. R. E. Wright** described two devices for inputs to computers. The first, the induction digitizer, was basically a transformer with a number of alternative flux paths. A common primary winding passed the alternative flux paths to a number of secondary windings and embraced flux, digits or windings. The polarity of the output voltage in respect of the primary voltage varied. Each digit of the output had a separate winding.

The second device was associated with digital differential analysers. In order to provide the incremental input on the channels of information, a 'synchro' could be used, and the sine-cosine output interpreted as a fourth-stage signal. By reference to previous readings it was possible subsequently to determine whether the count had advanced or retired by one, and accumulate this in the integrated digital differential analyser and therefore accumulate shorter positions.

SESSION 3.—PERIPHERAL EQUIPMENT—II

SOME ASPECTS OF XEROGRAPHIC DATA PRINTERS

By K. G. Huntley and J. Hughes, B.Sc., Ph.D.

The basic element used in making a xerographic reproduction is a photo-conductive selenium layer which is sensitized by uniformly charging its surface. The layer is selectively discharged when exposed to an optical image, forming an electrostatic image. This latent electrostatic image is made visible by applying suitably-charged finely-divided pigmented powder. Because of its relative impermanence, this powder image is transferred electrostatically to a permanent support, e.g. a sheet of paper to which it is fixed by the application of heat. The surface of the photoconductive layer is then cleaned of excess powder and recycled.

The re-usable character of the sensitive material in xerography makes it particularly suitable for use in an automatic machine. In the xerographic continuous printer the photo-conductive surface is cylindrical and the charging, exposure slit, development, transfer and cleaning positions are suitably spaced around the periphery of the drum. Such a printer, in conjunction with a device for displaying characters (which here will be restricted to some form of cathode-ray tube), can be used as a computer output printer.

The relatively low speed of present xerographic drums and a spectral response confined to the near ultra-violet and blue imposes certain limitations on the characteristics of the cathode-ray-tube screen. The general requirement is for a high-definition tube with a phosphor giving the maximum light output in the correct spectral range consistent with fast build-up time and short afterglow. Methods of displaying characters on a cathode-ray tube can be conveniently divided into schemes using special tubes and those involving conventional ones. In the special tubes the cross-section of the beam is made the shape of the character to be displayed, in one type by using a limiting stencil diaphragm.* Features of this method of display are:

(a) No complicated character-generating circuits are required.
(b) Comparatively low beam-current density.
(c) The character style is fixed.
(d) The character size is not electronically controllable.
(e) The electron-optical design of the tube is difficult.
(f) Replacement cost is high.

Schemes of character display using conventional cathode-ray tubes may be subdivided into those using a regular scanning raster, the character being displayed by the application of the appropriate brightness modulation waveforms, and those in which x- and y-deflection waveforms are generated appropriate to the required character (the character may be traced as either a continuous line or a succession of dots).

The first scheme is analogous to a television link. The fount of characters may, for example, be a matrix transparency scanned by a flying-spot scanner or a monoscope. Further systems include one in which a core matrix is used to store and generate the brightness-modulating waveform.† When a conventional cathode-ray tube is used the character-raster waveforms and the locating or 'addressing' potentials must be mixed electrically. This requires considerable scanning power, the amplifying circuits needing a large bandwidth and dynamic range. This may be overcome by using a tube with two independent means of scanning. One deflecting amplifier will provide the small deflections at high frequency required for the character raster, while a second will provide the large deflections, at a

Mr. Huntley and Dr. Hughes are with Rank Precision Industries, Ltd.

* Peterson, R. M., and Ritchart, R. C.: 'Recent Developments in Shaped Beam and Recording Techniques', I.R.E. National Convention Record, March, 1958, Part 3, p. 21.

† Wang, A.: 'High-Speed Number Generator Uses Magnetic Memory Matrices', *Electronics*, May, 1953, p. 200.

much reduced bandwidth, for addressing. In one such tube, electrostatic deflection is used for character forming and electromagnetic deflection for addressing.

Features of methods in this category are:

(*g*) The character storage and generation circuits are fairly complex.
(*h*) Electronic control of character size is easy.
(*i*) Very high peak current density is required, because of the inefficient use of the scanning raster.
(*j*) The type of character is easily changed.
(*k*) It is wasteful in time, since the beam is switched off for a large proportion of display period.
(*l*) The tube replacement cost is low.

In the second scheme suggested above, function generators produce the x- and y-deflecting waveforms appropriate to the character selected, and a blanking waveform is applied to the tube modulator to suppress the beam except during actual writing periods. Suitable continuous waveforms may be generated by methods similar to those proposed for analogue computers. Alternatively, the required waveforms may be quantized and generated as successive discrete levels, the display then being in the form of adjacent, or preferably slightly overlapping, dots. In one such system the dot position co-ordinates of each character are stored digitally in a core matrix and subsequently converted to deflecting potentials in a digital-analogue convertor.* Other methods include the use of passive networks where each level of the quantized waveform is represented by one element of the network. Methods in this category exhibit features (*g*), (*h*), (*j*) and (*l*), as before, but in this case the beam is on practically continuously during character display, and so the peak beam current-density is considerably less than before, although it is greater than in special cathode-ray tubes.

* British Patent No. 781340.

A printer using a method in the second category has been successfully demonstrated by the authors in which characters are displayed at the rate of 5 000 per second. The characters were formed by the successive display of up to 35 dots, each of duration 5 microsec. Adequate exposure was obtained with a beam current of 150 μA and an optical system with aperture $f/4{\cdot}5$, at a magnification of $1{\cdot}1 : 1$.

Some means for addressing the characters displayed across the screen must clearly be provided, and in the equipment constructed by the authors the following scheme is used. A 7-stage binary register followed by a digital-analogue convertor produces the horizontal deflection potentials. For normal sequential display the contents of the register are increased by unity on the completion of a character; in addition, they may be set to any desired number by means of orders from the input medium, so that the printing position may be indexed for tabulating purposes.

Of the two registers for vertical addressing, the first is cleared to zero at the commencement of each form and, by means of a count input, records the line number which is available under the printing head at any given moment. The contents of the second register are set up to the line number which it is next desired to print. The contents of the two vertical registers are compared in coincidence logical circuits so arranged that further reading of data from the input medium is inhibited when their contents are not equal.

A HIGH-SPEED LINE PRINTER USING OPTICAL METHODS

By G. G. Scarrott, Associate Member, and J. A. Freer, M.A., Graduate.

As computers become faster it becomes increasingly difficult to make mechanical printers to match them in speed of handling information. Yet in order to get most value from a computer, particularly in a data-processing system, it is important that a sufficiently fast output printer should be available. The printer to be described is an attempt to satisfy this need. An optical method has been adopted which allows the complete elimination of reciprocating mechanisms, so that the mechanical design can be very simple. A second feature of the design is that complicated electronic components and precise analogue-control waveforms have been avoided.

The essence of the method of producing a line of print is that master characters are illuminated as required and their images are projected by a lens on to a moving photographic recording medium. The master characters are opaque on the surface of a continuously rotating transparent cylinder. Positions are defined along the axis of this cylinder to correspond with each character position within a printed line. In all these positions the full set of available characters is inscribed round the circumference of the cylinder.

One of the limits to the speed of a photographic device is the exposure required by the sensitive material. Because the master characters are moving, it would normally be possible to allow only very short exposure or a blurred image would result. To compensate for this motion and to maximize the exposure, concave cylindrical surfaces are cut in the drum. Their curvature is such that when a character is projected it is apparently on the axis of rotation and therefore stationary.

Mr. Scarrott and Mr. Freer are with Ferranti, Ltd.

Simple unfocused cathode-ray tubes are used as sources of light. One tube is provided for every character position within a line, and a system of light guides directs the light from the face of each tube to illuminate one master character position. A ferrite-core matrix controls the sequence of turning on light sources. The rectangular matrix has a row for each of the available characters and a column for each character position in a line.

Data specifying one line of print is decoded into the rows, characters going to successive columns. Each column then contains only one core in the 'set' state. The line of print is formed during one revolution of the master cylinder. As each row of characters on the cylinder comes opposite the row of light guides, a full 'reset' current is passed through the corresponding row of the matrix. In being reset, any core which was previously in the set state emits a voltage pulse. This pulse is used to turn on the light source in the corresponding column for an accurately defined time. In this way the line of print is built up in the order in which characters pass the light sources, no matter in what columns these characters occur.

Since different characters are exposed at different times, it is necessary for the recording medium to be effectively stationary during the printing operation. If the medium were to be literally stationary, an intermittent motion would be necessary. To avoid this, a square prism is used to deflect the image formed by the projection lens. The prism rotates in synchronism with the master drum and causes an image displacement proportional to its angular rotation. Thus the photosensitive material may move at constant velocity, while the displacement of images ensures that all characters are, in fact, printed in a straight line.

A HIGH-SPEED STYLUS PRINTING SYSTEM SUITABLE FOR USE WITH MAGNETIC TAPE OR PUNCHED CARDS

By J. H. LUCAS, B.Sc.(Eng.), Member.

The stylus printing technique described offers the facility of defining the shape of a character by a train of electrical impulses instead of by the actual configuration of the character on the face of a type-wheel or type-bar. This enables characters to be selected by a switching operation, which can be either electro-mechanical or electronic, instead of by the physical movement of the type carrier. The combination of stylus printing and electronic switching offers the prospect of providing a reliable high-speed line-printer capable of meeting the requirements of high-speed data-processing systems, particularly for commercial applications.

The system described was developed with the primary object of providing a high-speed line-printer suitable for printing directly from information recorded on punched cards. A second stage of development has been directed towards providing a line-printer capable of accepting information serially at character rates compatible with magnetic tape.

The printing system is based on the use of a single stylus per character, and has not previously been described in technical literature. In the simplest form, the stylus consists of a thin flexible steel wire, the point of which is made to sweep to and fro across the surface of the paper, which is fed slowly forward at the same time. In this way the stylus point can be made to sweep a raster over the area to be occupied by a character. By interposing carbon paper or ribbon between the stylus and the paper, and by driving down the stylus at predetermined intervals, a pattern of dots can be printed which gives an acceptable representation of the desired character. A line of print is obtained from a row of styluses operating in parallel and having a common reciprocating drive to provide the sweep. Driving impulses are applied to each stylus through an electro-mechanical transducer, and can thus be generated as electrical impulses, which can readily be switched to the required styluses.

In the punched-card printer the pulse patterns for each different character are generated by commutator-type switches with segments appropriately spaced, which rotate once per line o rint. One such commutator switch is required per character. The styluses are driven through Bowden cables by moving-coil units similar in pattern to those used for loudspeakers. Connections between the commutator switches and the moving-coil driving units are effected through a relay switching matrix. Information punched in the card is sensed in parallel and used to set up the relay matrix during the interval preceding each line of print. Machines working on this system at 300 lines/min are in existence.

Such a system cannot be adapted to accept digital information in serial form, as required for use with magnetic tape, without the intervention of a separate buffer store to provide the required parallel setting. To avoid this necessity, an electronic type of switching matrix has been developed, which makes use of sub-miniature gasfilled triodes. These tubes combine the functions of storage and switching, as can relays, but they can be set (triggered) very much more rapidly. This property makes serial setting of the matrix a practical possibility, and avoids the need for additional buffer storage. Subminiature gasfilled tubes cannot switch at such high power levels as relays, and for this reason a power amplifier must be provided for each stylus coil-unit.

A system based on gas-tube matrix switching (using type GTR 120 tubes) has proved capable experimentally of accepting information serially at a character rate of well over 20 kc/s. The lower limit is set by the number of characters to be entered in the time available between successive lines of print, and is about 2 kc/s in a practical machine.

The lower power level at which switching takes place in the gas-tube matrix opens the way to more refined methods of generating the character pulse trains. A photo-electric system using perforated brass discs in conjunction with semiconductor photocells has proved quite satisfactory. The elimination of mechanical contacts both for character generation and for matrix switching has obvious advantages in increased reliability and reduced maintenance.

The printing speed of the system described above is limited to a practical upper limit of 300 lines/min by the response time of the moving-coil elements and by the required rate of lateral reciprocation of the styluses. Further developments are in progress to provide a type of stylus drive with a faster response and with the additional advantage of requiring a lower energy pulse for actuation.

The possibility of using ferrite cores as matrix elements instead of gasfilled tubes was originally rejected, owing to the difficulty of non-destructive read-out. It has now been shown that the need for non-destructive read-out can be avoided by storing the dot-position information in the matrix. The development of this idea is the subject of a separate contribution to the discussion.*

Mr. Lucas is at the Whyteleafe Research Establishment of International Computers and Tabulators, Ltd.

* ELLSON, A. H.: See below.

A FERRITE-CORE BUFFER STORE AND CHARACTER GENERATOR FOR USE WITH A HIGH-SPEED STYLUS PRINTER

By A. H. ELLSON, B.Sc.(Eng.), Associate Member.

The techniques which are described here have been developed with particular reference to the high-speed printing system which is the subject of an earlier contribution.* If such a printing system is to be connected directly to a high-speed computer, it must be able to accept the information for a line of print at the computer digit rate. Furthermore, the printer should operate 'on demand', i.e. it should be possible to control the paper-positioning and character-generating systems from the computer in such a way that, when the 'print' order is given, printing and paper feeding can commence with the minimum of delay.

A ferrite-core matrix is a compact and convenient buffer store which will accept information at rates up to about one complete character every 5 microsec. If the wiring of the matrix is arranged so that the required dot pattern is generated and stored as each character is read in from the computer, the normal

* LUCAS, J. H.: See above.

Mr. Ellson is at the Whyteleafe Research Establishment of International Computers and Tabulators, Ltd.

rotating-disc type of pattern generator can be eliminated. This facilitates the achievement of an 'on demand' printer operation with starting times of the order of 10 millisec.

A buffer store has been constructed which will accept 6-bit parallel-coded character information from a computer at a digit rate of 200 kc/s. Storage is provided for 96 characters in a line of print.

The store matrix consists of 96 columns (one for each character position) and 121 rows (one for every possible position of a dot on the scanning raster). There is a 2 mm ferrite core at the intersection of each row and column. When information is being inserted, each 6-bit group is decoded by a core-diode network to feed a half-current into one of the 64 possible paths entering the store matrix in the row direction. The decoder operates on the 'current-steering' principle first described by Karnaugh.* A set of six binary pairs of cores are switched by the input code pulses during a single digit time (about 2 microsec). Application of a resetting m.m.f. to these cores during a subsequent $2\frac{1}{2}$ microsec period induces voltages in certain of the windings which bias associated diodes in such a way as to steer an applied current pulse into one path out of the possible 64. Each path from the decoder corresponds to a particular letter or figure and is wired in the matrix to traverse only those rows which correspond to dot positions included in that particular character shape. Since each dot position may be common to several different characters, the matrix is threaded initially with 18 wires through each row of cores. (Experience has shown this number to be sufficient.) When the interconnections are made to form the character shapes, one of these wires is used as part of the input current path for every letter or figure which shares a particular dot.

In the column direction an additional half-current from a core-diode 'stepping circuit', driven in time with the input source, ensures that character positions are filled sequentially. Current-steering principles are again used for this stepping circuit.

When all 96 columns have been filled a second 'stepping circuit', driven by the printer mechanism, steers a current through each of the 121 rows in turn, at intervals corresponding to the dot positions in the scanning raster. The current in this case is made sufficient to reverse the state of any cores which were set during input, so that an e.m.f. is induced at any given dot time in every column wire which threads a set core.

Amplifiers connected to each of the columns are used to actuate the stylus drive. In this way patterns of dots are impressed on the paper in each character position to reproduce simultaneously the 96 characters stored in the line.

In this particular application the time taken to transfer 96 characters from the computer to the buffer is 1 millisec. The printer takes 55 millisec to print the complete line and 45 millisec to move to the next line. The arrangement is such that, if in the intervening 45 millisec the buffer has not been filled, the paper will stop. The scanning system continues to run and to send interrogation pulses to the core store. The store 'read-out' system, however, is not permitted to operate until a signal is received indicating that the buffer has been filled. When this occurs, 10 millisec is allowed for the paper to accelerate to speed and printing begins with the first dot of the next raster line.

The larger number of matrix elements required to store the complete dot patterns for a line of print is compensated for by the elimination of a separate character-generating device. The use of a ferrite-core matrix to perform both functions leads to a compact, non-synchronous and reasonably economical system particularly suitable for operating the printer direct from a computer.

* Karnaugh, M.: 'Pulse Switching Circuits using Magnetic Cores', *Proceedings of the Institute of Radio Engineers*, 1955, 43, p. 570.

A HIGH-SPEED SERIAL PRINTER

By F. R. Thomas, Associate Member.

The increasing use of electronic computers for commercial applications has resulted in the need for high-speed output printers capable of producing documents of the type customary in business transactions. In some of these applications the volume of printing required makes the use of high-speed line printers essential, and various mechanical and electrographic printers are now available to meet the needs of this particular field. The nature of these machines and the need for the storage of a complete line of information for each printing cycle necessitate a primary cost which is too high to permit their economical use in installations where such a high rate of output is not essential. For such types of installation, which include small data-processing systems and heavily loaded scientific computing systems, the requirements are likely to be met by a printer capable of medium-speed operation and automatic layout control, but requiring only simple controlling apparatus.

The availability of such a machine would also confer economically a remarkable degree of flexibility in regard to

(*a*) Overall speed, since they could be used in groups.

(*b*) The capability of dealing simultaneously with different documents, since they can be separately programmed.

(*c*) The maintenance of continuity of service in case of breakdown, owing to the possibility of sharing the computer output between several relatively low-priced units.

Mr. Thomas is with Creed and Co., Ltd.

Consideration of the various design possibilities has led to the conclusion that a mechanical serial (character-by-character) printer would provide the most effective solution. This solution makes for simplicity in the driving circuits, since only modest buffer arrangements are needed and control from an intermediate tape store can be effected very simply. Taking into account the related factors of speed, maintenance and prime cost, it is believed that a printing rate of 100 characters/sec will be appropriate for such a printer. The net speed would be somewhat lower, since the functions associated with the movement of the type carriage to the beginning of a new line and to tabular positions occupy a certain amount of time, although this can be minimized by suitable design. A suitable method for obtaining such a speed, which is roughly ten times higher than that of current serial printers, is the use of mosaic printing, in which the characters are not selected from a fount of solid type pads but are composed by the use of a 5×5 matrix of printing styluses and printed through a one-time carbon strip on to the paper. Sufficient power can be made available in the movement of the styluses to permit up to five carbon copies to be produced.

The machine can be internally programmed to provide, not only horizontal tabulation, but also vertical tabulation or form throw.

DISCUSSION ON
PERIPHERAL EQUIPMENT—II

The discussion was opened by **Mr. E. A. Newman,** who asked all contributors for an indication of the capital cost per digit for the various printers which they had described. All of the speakers regretted their inability to provide this information on the grounds either that the machines described were still in the development stage or—in the case of Mr. K. G. Huntley—that he considered it unethical to quote costs, although he did feel that the xerographic printer would compare very favourably with slower printers in terms of cost per character per second.

Mr. Newman then queried the feasibility of a printer which combined techniques taken from more than one of those described. Mr. Huntley referred to common ground existing between the techniques incorporated in the xerographic machine and the optical printer described by Messrs. Scarrott and Freer.

Dr. J. M. M. Pinkerton described the advantages of on-line printers, but indicated the necessity for fast paper feed with this type of operation.

Mr. K. G. Huntley, in reply, indicated that the problems of on-line printing could be solved using magnetic tape as a direct output, which, in turn, could drive the printers off-line in their own independent time schemes. An alternative would be to use several printers on a time-sharing basis. Mr. Huntley further indicated the merits of continuous paper feed, and contended that the high speed of operation of the xerographic printer outweighed any disadvantage due to such techniques.

Mr. J. A. Freer agreed with Dr. Pinkerton in the desirability of on-line printers, and suggested that it might be possible to remove some of the disadvantages of continuously feeding paper by recording on microfilm.

Mr. J. H. Lucas indicated that servo systems for paper-feed control had been considered, but did not appear to have a decisive advantage over a straightforward high-precision mechanical system, and at the same time were expensive.

Mr. A. H. Ellson pointed out that high-speed paper-feed mechanisms did not start and stop merely the paper. The problem of starting and stopping paper for simple line feed was not difficult. Even the problems associated with considerable throws of paper were not incapable of solution.

Mr. G. D. Royle felt that the session was devoted to the description of two categories of printer, that using reciprocating mechanisms and that using plain rotating principles, the latter being capable of fundamentally higher speeds. The applications of high-speed printing appeared to be threefold. The first was to produce information which acted as input to further machines, which appeared to be an unlikely possibility. The second was the use of high-speed printers to provide short bursts of printing which would not hold up the computer. Finally, there was the production of large volumes of printing with the corollary of a large number of human readers. So far as the use of preprinted stationery was concerned, the xerographic and optical printers made this unnecessary, unless printing on the backs of forms o printing in different colour was required.

Finally, Mr. Royal commented on the main weakness wit ultra-high-speed printers lying with an inability to forı throw at high speed, thus effectively reducing the net speed o operation.

Dr. K. Chandler felt that both the xerographic and optic printers described would be expensive, and asked whether ar reduction in speed might result in lower costs, while still pro viding adequate speeds for most purposes. **Mr. K. G. Huntle** replied to the effect that, despite a small economy resultir from speed reduction, the cost per character per second woul certainly increase.

Dr. Chandler went on to illustrate a means whereby the spee of the optical printer might be easily increased. **Mr. G. C Scarrott,** in reply, felt that Dr. Chandler's suggestion would, i effect, result in the printing on the paper of a series of ve long rectangles, with the effect that the available light woul need to be spread over a much wider area with the resultir disadvantages.

Mr. A. D. V. Ridlington queried the length of life of th selenium surface of the xerograph drum, and **Mr. K. G. Huntle** replied that this was approximately 120 000 ft of paper.

Mr. N. A. F. Williams asked whether any hydraulic resonanc had been experienced on the high-speed serial printer, particular at different printing speeds. **Mr. F. R. Thomas** indicated that r such troubles had been experienced, particularly because th speed of printing remained constant.

Mr. B. W. Pollard proffered the suggestion that pre-cu stationery could be fed into the xerographic printer a sheet at time and printed as it came out. Following on this, he querie the possibility of writing individual lines on the selenium rolle as the information became available from the computer, ar printing off the whole contents of the roller when filled or whe the printing operation was called for. **Mr. K. G. Huntley** replie to the effect that this represented a practicable proposition.

Mr. L. R. Crawley referred to the necessity with high-spee paper transport of printing the form outline at the same tin as the information. He also requested details of the work bein done to provide colour copies on both the xerographic ar photographic equipment. **Mr. K. G. Huntley** indicated th work was proceeding on the question of colour, but at prese nothing was available.

Finally, **Mr. J. K. Webb** queried the relative merits of th serial-mode matrix printer against a line-at-a-time matrix printe He also referred to the improved appearance derived from 7×5 matrix as opposed to 5×5 one. **Mr. F. R. Thom** replied that both the serial mode of operation and the choi of a 5×5 matrix were dictated by an endeavour to achie simplicity and economy.

SESSION 4.—LOW-TEMPERATURE STORAGE AND SWITCHING DEVICES

INTRODUCTORY LECTURE

By D. R. Young, Ph.D.

In 1956 Buck* published a paper describing the development of a wire-wound switching device called a 'cryotron', which used magnetic control of the state of a superconducting wire. He also described several circuits of interest for application to computers. The speed of operation of these circuits was in the kilocycles-per-second, range, owing to the L/R time-constants of these devices.

In 1957 Crowe† and Garwin‡ described the operation of a memory device utilizing thin superconducting lead films which operated in the millimicrosecond region, thus demonstrating the potential high speed of operation of superconducting devices.

To provide the switching functions needed for a high-speed low-temperature computer it has been necessary to develop a high-speed cryotron. This has also utilized thin superconducting films and has been named the 'planar cryotron'.

Successful circuit operation requires that a cryotron is able to control a larger current than is needed to do the controlling. A significant portion of the work has been directed with this requirement in mind, and the resulting techniques used for the solution will be described.

The progress made to date in film reproducibility will be indicated. This will include data on the effect of the gas content of tantalum on its superconducting properties, which indicates the importance of the degree of vacuum obtained during the evaporation process.

The operation of some simple circuits will be described and a graphical method for predicting the circuit performance from the control characteristics of the cryotrons used will be presented.

* Buck, D. A.: *Proceedings of the Institute of Radio Engineers*, 1956, 44, p. 482.
† Crowe, J. W.: *I.B.M. Journal of Research and Development*, 1957, 1, p. 304.
‡ Garwin, R. L.: *ibid.*

Dr. Young is at the I.B.M. Research Center, New York.

THE SUPERCONDUCTION OF SWITCHING AND STORAGE

By K. A. G. Mendelssohn, M.A., D.Phil., F.R.S.

Superconductivity was discovered in 1911 by Kamerlingh Onnes; but, in spite of the lapse of almost half a century, we still have no completely satisfactory theory of the phenomenon. On the other hand, an enormous number of experiments have been performed with superconductors, so that we now possess a well-defined phenomenological picture of the unusual effects involved. The salient feature is, of course, the complete loss of resistivity, which has been demonstrated by maintaining a persistent current in a simple ring for over a year without measurable decay. The disappearance of resistivity occurs, for a pure metal, discontinuously at a given temperature, the transition point, T_c, being characteristic of the metal and ranging from 0·5° K to 10° K for pure metals. The highest T_c so far has been found in a compound of niobium and tin (about 18° K), and for a number of reasons it appears at present unlikely that appreciably higher values of T_c can be obtained. This means that devices embodying superconductors require the use of liquid helium. (Liquid hydrogen under much reduced pressure, which could be used in conjunction with niobium compounds, is hardly a feasible proposition.) With adequate quantities of helium gas from natural wells, and with commercially produced helium liquefiers—both of which are now available—the temperature region below 10° K is open to practical applications. The least complication is encountered by working at 4·2° K (the boiling point of helium), or just above it. The critical point is at 5·2° K, and temperatures down to about 1° K can be obtained by pumping off the vapour.

Only certain metals, none of them monovalent, have been found superconductive. Of these, only those whose T_c is not too low are, at the present state of cryogenics, suitable for application in switching devices. They are: niobium (9·2° K), lead (7·2° K), vanadium (5·1° K), tantalum (4·5° K), lanthanum (4·4° K), mercury (4·1° K), tin (3·7° K), indium (3·4° K) and thallium (2·4° K). At our present state of knowledge the use of alloys or compounds is inadvisable, and lanthanum, mercury and thallium have undesirable properties. For reasons stated below, niobium, vanadium and tantalum present difficulties, which leaves lead, tin and indium, of which only lead has a T_c value above 4·2° K.

Superconductivity can be quenched by a current or by an external magnetic field corresponding to it. The critical data are usually expressed as temperature-dependent fields, H_c, which are characteristic of the metal. For all superconductors $H_c = 0$ at T_c and becomes temperature independent near absolute zero, the H_c/T_c curve being roughly parabolic in shape. The value $H_{c(T=0)}/T_c$ is of the order of 100 oersteds/deg for the soft metals and about 200 oersteds/deg for the hard ones. In addition to being perfectly conducting, the superconductors are also perfectly diamagnetic, and these two conditions are not consequences of each other in Maxwell electrodynamics. Complete expulsion of magnetic flux is strongly dependent on physical and chemical purity, and for this reason the hard metals are at present more difficult to use than the soft ones. This is a drawback with niobium, vanadium and tantalum, which would otherwise be admirable for switching devices.

Superconductive switching devices generally work through the quenching of superconductivity by an external field, usually produced by another superconductor. The transition between the superconductive and normal states is attended by thermal and electrodynamic effects, which tend to reduce the speed of switching. Both of these are volume effects, and it is therefore necessary to use fine wires or thin films for fast switching. For the above-mentioned reason, most experiments in the latter field have been made with lead, and switching speeds of 25 millimicrosec have been reported. Circuits are usually so arranged that the switch is open in the normal state and closed in the superconductive state. Moreover, the use of persistent currents allows the construction of excellent memory devices for large storage devices to which access can be rapid. If to these features we add compactness and negligible power dissipation, the advantages may easily outweigh the use of liquid helium.

Dr. Mendelssohn is Reader in Physics at the University of Oxford.

SUPERCONDUCTING STORAGE DEVICES

By J. M. Lock, Ph.D.

In a superconducting circuit it is possible to set up a persistent current which, because the circuit has zero resistance, will continue to flow indefinitely, and left- and right-handed persistent currents can then be used as the two states of an information-storage device. To switch the device from one state to the other, use is made of the fact that, if the current in the circuit is raised above a certain critical value, superconductivity is destroyed, the normal resistance returns and the current decays. Such changes can be brought about by means of a drive wire linked inductively with the superconducting circuit, and positive and negative drive-current pulses then produce persistent currents in opposite directions.

The speed with which such a device can be switched from one state to the other depends, among other things, on the rate at which the supercurrent decays when the normal resistance is restored; in other words, on the L/R time-constant of the circuit in its normal state. To make this time-constant very small, the storage elements are made in the form of a thin film, prepared by evaporation of the metal *in vacuo*, and each element is only a few millimetres across. L is then small and R reasonabl large, so that L/R can be made less than about 10 millimicrosec The switching speed also depends on the rate at which the hea produced when the superconductor is driven resistive can b dissipated, and this in turn depends on the nature of the substrat on which the film is deposited.

The behaviour of thin films of lead and tin has been studie at temperatures down to $1 \cdot 5°$ K, and using tin—which has a lower heat capacity and hence tends to have a lower therma time-constant than lead—persistent-current elements have bee made which can be switched in less than 50 millimicrosec. I addition, critical current measurements have been made on bot planar and cylindrical evaporated films, and have yielded infor mation on the L/R and thermal time-constants of the storag elements. They also lead to an estimate of not less than 20 Mc/ for the maximum rate for repetitive switching.

It is considered that it would be feasible to use this type o storage element as the basis of a large computer memory con taining as many as 10^7 bits, and with a random access time o less than 50 microsec.

Dr. Lock is at the Royal Radar Establishment.

THERMAL RELAXATION TIMES IN THE TRAPPED-FLUX STORAGE DEVICE

By O. Simpson, M.A., Ph.D.

In any storage or switching device employing a superconductor the cycle of operations will necessarily include an interval when current is flowing in normal metal. The Joule heating accompanying the normal current causes a temperature rise in at least a part of the structure, and the thermal relaxation time consequently imposes a limit on the maximum repetition frequency of the device.

Experiments have been performed at the Services Electronics Research Laboratory to study the thermal relaxation of superconducting lead films on various substrates. In order to apply the results to a case of practical importance, the films were made as trapped-flux storage elements, or Crowe cells. After measurements had been made of their storage characteristics, the cells were modified by introducing an open-circuit in the persistent current loop; this allowed further measurements to be made on the crossbar alone. The results are, however, qualitatively applicable also to thin-film cryotrons, or any other device employing a superconducting film deposited on an insulating substrate.

The experiment consists in applying in succession two current pulses to a parallel-sided strip of superconducting lead film immersed in liquid helium. The pulses have a half-width of 10^{-8} sec, and the time interval between them can be varied from zero to 10 microsec. For convenience of observation a repetition rate of a few kilocycles per second is used. The amplitude of the first pulse is just sufficient to drive the strip normal; this causes a temperature rise of about 3° K. Th amplitude of the second pulse, following after an interval t, i adjusted to the value $I_c(t)$ at which it is just sufficient to driv the film normal again. Since the critical current is a decreasin function of temperature, the amplitude of the second pulse wi be less than that of the first pulse if the film has not cooled t the temperature of the helium bath. The measurement i repeated over the relevant range of delay times t. If the critica current is also known as a function of temperature, the value of $I_c(t)$ can be used to deduce the cooling curve. The calibratio of the critical current, $I_c(T)$, is made on the same specimens in separate experiment.

Similar films were deposited directly on to three substrates– glass, mica and sapphire. The experiment was performed a $4 \cdot 2°$ K and also below the Λ-point in helium II. Preliminar results show the following:

(*a*) The average thermal relaxation time on glass is $1 \cdot 5 \times 10^{-7}$ se

(*b*) The relaxation times on mica and sapphire are 10–20 time shorter than those on glass.

(*c*) The relaxation time is not materially reduced in helium I showing that the heat is dissipated predominantly in the substrate.

(*d*) The maximum repetition rate of a film, working as a trappe flux storage cell, varies between 30 and 500×10^{-9} sec, and correlated with the thermal relaxation time.

The lead-film storage cells were made by Mr. R. Broom, wh also carried out the experimental work. Full details of th results will be included in a subsequent publication.

Dr. Simpson is at the Services Electronics Research Laboratory.

DISCUSSION ON LOW-TEMPERATURE STORAGE AND SWITCHING DEVICES

Dr. J. M. Lock opened the discussion by asking **Dr. D. R. Young** whether he or his associates had made any direct measurement of the switching speed of the thin-film cryotron. Dr. Lock's calculations indicated a switching time of 100 millimicrosec using Dr. Young's value of $L = 10^{-11}$ henry/cm. Dr. Young replied that, with a circuit length of 1 cm, L/R for their cryotron was calculated to be 3×10^{-10} sec, assuming a resistivity of 1 microhm-cm, which is higher than that observed with tin films. In practice, circuits have operated more slowly in the 1 Mc/s range with circuit lengths of about 5 cm; and he did not know whether they would get to the millimicrosecond range.

Mr. T. O. Stanley was pleased with the emphasis placed on thermal effects, because of their great engineering importance in the building of cryotrons. He asked whether the dependence of the thermal time-constant with respect to the thickness of the thin film had been measured. **Dr. J. M. Lock** replied obliquely, saying that there would be a longer thermal recovery time with thicker films, but the only reason for making a thicker film would be to increase the critical current (which could more easily be done by increasing the width of the film, thus increasing the area of contact and leaving the thermal time-constant much the same).

Mr. L. P. Morgan questioned Dr. Lock on the drive currents which he had used with his Crowe cells, and asked whether it were possible to make his units 100 times smaller, because of the very large path length which would be involved when wires threaded all bits of a 10^7-bit store. Dr. Lock replied that the critical drive current was dependent on the temperature of the elements, and he usually worked with drive currents of 0·2–0·5 amp. When a Crowe element is switched about LI_c^2 of power is dissipated—about 10^{-10} joule in practice. When building these elements into a large store he proposed to have planes of elements 8 in square, stacked with spacings of about ¼ mm. He believed that it was possible to attain a random-access store in which wires linked only a single row of elements, which meant that the length of wire involved would be only 8 in or so, i.e. a transit time of less than 1 millimicrosec. The necessary selection circuits would also be in the liquid-helium bath.

Dr. B. H. Rhoderick described an apparently simple experiment in which rectangular pulses, about 0·5 microsec long, were pulsed through a thin-film strip of lead and the voltage across it observed with respect to time. Until the amplitude of the current pulse exceeded the critical current, no voltage was observed, but above the critical current a steadily rising voltage was observed after an initial time delay. When the voltage pulse was long enough to climb to a value corresponding to the normal resistance of the lead strip at a temperature just above the transition point, the curve broke and the voltage stayed almost constant at that value. It is known that such lead strips will switch in a Crowe element in about 10 millimicrosec. Therefore, only about a tenth or a fiftieth of the normal resistance returns during the drive-current cycle. Secondly, it is deduced that when making a planar cryotron a larger critical field has to be used for dynamic operation than that measured under d.c. conditions, and any measurement of current gain with direct current may not be applicable with very short pulses. **Dr. D. R. Young** suggested that the slow part of the voltage rise might be due to the thermal spreading effect, wherein the switching process started at some imperfection in the film and spread across it. Others had mentioned that the critical currents measured for devices were considerably smaller than those expected, which was thought to be due to these imperfections in the film, and that slow switching times were found in these conditions. When the switching rates of the transition from the superconducting to the normal state in thin films were studied with an applied magnetic field rather than by self-current, the slow part of that voltage response did not appear and the critical fields observed compared more closely with those expected from bulk measurements. Dr. Rhoderick expressed disagreement with the thermal spreading explanation.

Mr. R. Hayes, who had been doing some work on the switching time of small bars of germanium at liquid-helium temperatures, asked for information about impact ionization phenomena, but could get no reply from those present.

Dr. J. E. R. Young asked about the crystalline nature of Dr. Simpson's sapphire substrate, and wondered whether any work had been done on epitaxial deposition of thin films on the faces of good crystals with a view to getting a better impedance match for thermal vibrations at the interface between metallic film and substrate. **Dr. O. Simpson** answered that he used cleaned and polished commercial sapphire and that he was very interested in the relative natures of the substrate and the film. Films had been examined with the electron microscope in the hope of correlating their structure with the nature of the substrate. No correlation had been found, but films which definitely reach superconductivity were seen in the electron microscope to have no obvious continuous path across them, and displayed a complicated structure of unconnected islands of metal. The crystalline structure of the lead films did not often coincide with these islands and much more work was needed in this area.

Dr. K. A. G. Mendelssohn mentioned measurements by Professor Blackman which indicated that the state of the metal film was more insensitive to the crystal structure of the substrate than might have been expected, or that other factors in depositioning the film seem to be more important.

Mr. T. O. Stanley asked about the effect of the temperature of the substrate during the evaporation of the metal film, and Dr. Mendelssohn mentioned detailed experiments done at Bristol in the early 1930's to indicate that the temperature of substrates affected the metal films deposited on them in a complicated way. Generally speaking, at very low temperatures the atoms being deposited had very low mobility and there was a likelihood of a somewhat amorphous film. At high temperatures there was much more tendency for the atoms to form a crystalline arrangement. A question which he thought should be asked was: how permanent is an element? With single crystals it was found that their conductivity and critical field were likely to change in the course of some months at room temperature. **Dr. O. Simpson** replied that lead cryotrons had not changed their properties over several months, even though taken in and out of a low-temperature bath many times.

In closing, **Mr. J. K. Webb** emphasized the commercial importance of cryogenic memories if they could be produced at a penny a bit.

SESSION 5.—SPECIAL ASPECTS OF LOGICAL DESIGN—I

INTRODUCTORY LECTURE

By M. V. Wilkes, M.A., Ph.D., F.R.S., Associate Member.

'Logical design' is the term applied to the overall design of a digital computer, or computing system, regarded as an assembly of switching and storage elements together with various items of peripheral equipment. 'Systems design' would perhaps be a better term, but 'logical design' is by now firmly established as the accepted usage.

The objects of good logical design are twofold: first, to secure the greatest economy in the equipment required to meet given requirements, or to obtain maximum speed from a given quantity of equipment; secondly, to meet as fully as possible the requirements of the user as regards speed, facilities provided and convenience in use. While responsibility for logical design is included among the other responsibilities of the design engineer, he will, if he is wise, consult someone who has had experience from the user's point of view. Indeed, some major contributions to logical design have been made by people who are not engineers. It has been noticeable for some time that certain trends in computer development are tending to divorce logical design to some extent from engineering design. One example is the provision of 'building blocks' which can be interconnected in any desired manner provided that certain loading rules are satisfied. Other developments having similar effects will be pointed out in the lecture.

Developments in the engineering design of basic components influence the economic aspects of logical design, and it is important to re-examine from time to time the extent to which formerly accepted principles are still valid. Two contexts in which this seems necessary at the present time relate to the provision of 'dead' or non-erasable storage, and to the importance which should be attached to the efficiency with which floating arithmetic operations are performed in machines intended for scientific calculation. The advent of time sharing has also raised some interesting points, particularly in relation to the relative economics of on-line and off-line equipment in business data-processing systems.

Dr. Wilkes is Director of the University Mathematical Laboratory, Cambridge.

PROGRAMME-CONTROLLED TIME SHARING

By C. Strachey.

It is generally accepted that a considerable reduction in the cost and complexity of the peripheral equipment for a large fast computer can be brought about by time sharing. This means that the main programme is automatically interrupted for a short period in order to allow the central computer to perform the logical and arithmetic operations required by the peripheral units. If the process of controlling the peripheral equipment is not to consume too large a portion of the total time available, the speed of the central computer must be very high in comparison with that of the peripheral equipment. This means that even the fastest computer projected at present, with operation time measured in a few microseconds, would have some difficulty in assembling words from a fast magnetic-tape unit with no buffer storage at all. If the word is assembled by the tape unit (which would require one shifting register and one single-word buffer) the requirements for the main computer are much less stringent.

If the time-sharing system is to be manageable from the aspect of the programmer, it is essential that the whole process should be automatic and that sufficient interlocks should be provided to avoid any possibility of confusion. This requirement inevitably introduces a considerable degree of complexity in the way in which the whole scheme operates. The best way to deal with a logically complicated situation is to use the power and flexibility of the main computer—in other words, to write a programme. It seems inevitable that the actual control of the peripheral equipment should be by a special programme which is more or less permanently stored in the main computer.

A more detailed investigation into the sort of programme required shows that the actual process of time sharing itself must sometimes be under programme control. This means that it must be possible to terminate a sequence of orders with a 'wait' instruction, which allows control to revert to a programme of lower priority. It may also be necessary for a programme (instead of a piece of peripheral equipment) to institute a sequence of orders of higher priority than itself.

It is evident that these fixed programmes can conveniently be held in some form of non-erasable store with a very rapid read-out time. It is also clear that they will need a certain amount of working space for storing numbers, counters, addresses, etc. These should probably be ordinary storage registers of the fastest kind which is economically justified. In order to avoid the complicated sort of error which could happen if a main programme altered these registers by mistake (owing to a programme fault), they should be accessible only from instructions in the fixed store.

When two or more programmes are using the main store on a time-sharing basis (a scheme which has much in its favour from the aspect of keeping a fast machine continuously in operation), it will be desirable to have interlocks on the main store in order to prevent one programme interfering with another. It would increase the flexibility of the machine if these programme interlocks were themselves under programme control. (They must necessarily be alterable, or it would never be possible to use the whole machine for a single large problem. There are obvious disadvantages in only altering them manually.) In order to avoid complications if one programme runs wild, it is desirable that these basic instructions should also only be available to the fixed store.

We thus get the concept of a general controlling programme or 'director', which is entirely contained in a fixed store, and as well as having access to all the ordinary store and ordinary instructions, has a private working store and an additional set of instructions which are not accessible to the normal programme. In a way the 'director' forms a sort of 'hyper-computer' which controls the operation of the main machine.

Mr. Strachey is with the National Research Development Corporation.

LOGICAL METHODS OF SPEEDING 'CARRY' IN PARALLEL COMPUTERS

By D. J. Wheeler, M.A., Ph.D.

A comparison is made between the different logical methods of arranging that 'carry' propagation time in parallel computers is not too long. Three main methods are discussed; first, 'carry' end detection, which relies on statistics and circuits to indicate the end of the propagation process; secondly, 'carry' forcing methods, which ensure that the sum is available in a short time, independently of the actual values of the numbers involved; and thirdly, 'carry' by-pass methods, which appear to show most economy.

Dr. Wheeler is at the University Mathematical Laboratory, Cambridge.

A FAST ADDER USING DIODE LOGIC

By G. Ord, M.Sc.

The adding circuits and associated registers have been designed with the aims of providing a fast adding cycle and, by the simple echnique of repetitive addition and shifting, a fast multipliation time. Since the shifting time for the register can be as hort as 200 millimicrosec, the adding circuits have been designed o have an operation time of the same order of magnitude.

In the design of adding circuits the primary consideration is o provide a 'carry' propagation path which allows a 'carry' to ropagate from the least to the most significant digit in the hortest time possible. Such a 'carry' occurs when the leastignificant digit of each of the incident numbers is '1' (thus proucing what is termed a first-order carry') and all the other igits of corresponding significance differ. In all stages of the dder, with the exception of the least significant, the 'carry' enerated is termed a 'second-order carry' and a gating circuit et by the non-equivalence of the incident numbers allows an nput 'carry' to be passed on to the next stage of each adder. The 'carry' propagation path thus consists of a series of gating ircuits of the 'and' type. However, with other configurations f incident numbers a first-order 'carry' may be generated in ny stage of the adder, and thus provision must be made for hese carries to be introduced into the 'carry' propagation path. This logical arrangement holds for many systems of addition. The basic logical circuit for each stage of the 'carry' path thus onsists of an 'and' gate which passes a second-order 'carry', followed by an 'or' gate which allows a first-order 'carry' to be introduced into the 'carry' path.

With low-hole-storage diodes in a diode logic system it was found that a negative-going voltage step of 3 volts and 60 millimicrosec duration provided by a low-impedance source was delayed by 15 millimicrosec. This was a shorter delay than had been expected from a calculation based on the currents used in the diode logic and stray capacitances involved. Measurements have indicated that hole storage in some of the diodes is providing a surge of current which assists in charging stray capacitance and removes holes stored in other diodes. After each stage of diode logic an emitter-follower is introduced, mainly to provide a low-impedance source for the next stage of diode logic. Restandardization of voltage level is necessary after three stages. With these additions, the average delay per stage is 20 millimicrosec.

As it stands the system is too slow. It had been intended to use a 'carry' detection technique which involves two 'carry' paths, but it was suggested in discussion with Dr. D. J. Wheeler that a 'skip' or 'jump over' technique might be more economical. In the 'skip' technique, groups of digits are examined to see whether a 'carry' can be allowed to skip over several stages. This technique has been introduced, and the 'carry' propagation path has skips over four stages. The average delay per stage is reduced to approximately 8 millimicrosec. When the skip technique is used it is necessary to disconnect the adding circuits from the incident numbers until an 'add' operation is needed.

Mr. Ord is at the Royal Radar Establishment.

DISCUSSION ON SPECIAL ASPECTS OF LOGICAL DESIGN—I

The discussion was opened by **Mr. C. Strachey,** who felt that he novel suggestion of **Dr. M. V. Wilkes**—that computers hould be designed for variable-accuracy working—might well hange the mathematical methods used in standard subroutines rom the second-order processes involving multiplication at resent used to digit-by-digit methods of extracting results.

Subsequently it was the contribution by Mr. Strachey which timulated most comment, and it was evident that 'interrupt' rrangements in machine design were attracting much thought n computer circles. **Mr. J. C. Gladman** outlined a scheme on vhich he had been working which approached the problem sing more equipment than the programmed method which Mr. trachey described. Mr. Gladman's method involved a shiftegister type one-word buffer store and a counter register assoiated with each piece of peripheral equipment (a magnetic-tape nit, for example). In addition, equipment associated with the nain-store control detected when a complete word had been assembled in the buffer store and so was ready for transfer to the main store. At this time the main programme was interrupted to perform this transfer. The main store could be used in this way during a multiplication or division order, for example, so that the computer can make more efficient use of its store.

Mr. D. B. G. Edwards felt that there might be economic difficulties in building a computer with a large enough store to allow for two main programmes to run in parallel. **Mr. C. Strachey** thought the size of stores contemplated for machines in this country had for too long been too small and that computer designers should plan for stores of between 16 000 and 100 000 words, all in the same medium. Such stores were essential if automatic programming techniques were to be adopted, and such techniques were essential.

In reply to a question from **Mr. G. C. Tootill, Mr. C. Strachey** showed that the difference between a conventional programmed

subroutine and a programmed time-sharing system was that the entry to the conventional subroutine is under normal programme control but that the entry to a subroutine associated with a time-sharing system is forced by computer control unit. It was stressed by Mr. Strachey, in reply to a point raised by **Mr. N. A. F. Williams,** that he considered it was essential to supply equipment in a computer to stop two 'parallel' programmes interacting one with the other. The aim should always be to make the job of the programmer as simple as possible.

Regarding the question of fast adding circuits which several speakers had described, some discussion ensued on the relative merits of synchronous and asynchronous techniques. A point made by **Dr. D. J. Wheeler** in reply was that even with synchronous systems it does not follow that an addition, for exampl need take a fixed time, since 'end of carry' detection can termina the process.

Discussing a point from **Dr. M. V. Wilkes's** lecture, **Dr. C. I Lindsey** questioned whether it was necessary to scale the larg factor in a floating-point multiplication operation, since required equipment to determine which factor was the larg and similar results could be obtained by scaling either facto **Dr. M. V. Wilkes** replied that one might have to compromise c this matter, but that the finding of the smaller factor was not s very difficult, because with non-standardized numbers it wa necessary only to see which had fewer zeros before the fir significant digit.

SESSION 6.—SPECIAL ASPECTS OF LOGICAL DESIGN—II

A SYSTEM OF CONTROL FOR A FAST COMPUTER

By G. Ord, M.Sc.

When a problem is introduced into a digital computer, the programme to carry out the problem consists of instructions from the instruction code of the computer. Each instruction may call for a series of actions, e.g. the opening of a gate between two registers of the arithmetic section followed by a shift in significance of the contents of one of the registers. In early computers the individual instructions required a few separate actions to be carried out sequentially, and it was possible to design separate electronic circuits to provide a sequence as well as the actions needed for each instruction. In more recent computers the individual instructions have increased considerably in complication, some requiring up to 50 steps in sequence. It is possible to design a sequencer for each of the instructions as before, but obviously it is preferable to design one apparatus which can be arranged to provide all the sequences required. Wilkes, Wheeler and Renwick* have described such an apparatus. Although in principle the system is applicable at high speeds, the use of the ferrite cores currently available limits the speed of the system to approximately 500 kc/s.

In the design of the proposed sequencer advantage has been taken of two facts, namely

(*a*) All the shifting, counting, transferring and exchanging operations in and between registers take the same time (about 200 millimicrosec) and comprise the large majority of the actions needed in the arithmetic section of the computer.

(*b*) In four out of five cases the detailed steps of an instructic follow one another in a regular sequence.

A shifting register is arranged to provide an active state only one of its outputs, all the other outputs being inactiv Unless prevented from doing so, the active state progress regularly from one digit of the register to the next eve 250 millimicrosec. To the output of each digit is connected set of transistor switches controlled by the instruction. A instruction controls the same number of switches as there a steps in its sequence, but only one of the switches connected any digit output is operated at the same time. The outputs the switches are connected to circuits which, when energize provide gating pulses to various parts of the computer. gating-pulse generator is therefore energized only when appropriately timed output from the shifting register is route to it by a switch.

If an alteration to the regular sequence is required, it arranged that at a particular time the position of the acti state in the register is altered. For some actions which requi more than the normal sequencing time the regular progressic of the active state along the register is inhibited.

Since the transistor switches are set up immediately after t instruction is known and are not altered during the sequen corresponding to that instruction, the time needed to set t switches can be as long as 0·5 microsec without affecti seriously the overall speed of the sequencer. The sequenc speed is determined primarily by the speed of the shifti operations, which can be as rapid as 4 or 5 Mc/s.

* Wilkes, M. V., Renwick, W., and Wheeler, D. J.: 'The Design of a Control Unit of an Electronic Digital Computer', *Proceedings I.E.E.*, Paper No. 2365 M, June, 1957 (105 B, p. 121).

Mr. Ord is at the Royal Radar Establishment.

PARALLEL ADDITION IN DIGITAL COMPUTERS: A NEW FAST 'CARRY' CIRCUIT

By T. Kilburn, M.A., D.Sc., Ph.D., Member, D. B. G. Edwards, Ph.D., M.Sc., Associate Member, and D. Aspinall, M.Sc.

When two numbers X and Y are added together, the kth digit of the answer is dependent on the kth digits in X and Y and also on the 'carry' digit which could be initiated by X and Y digits of less significance. If addition is carried out serially, the least-significant digits of the numbers are added first and it is necessary to delay any 'carry' indication until the next more-significant digits are processed. To add two n-digit numbers in this way takes n digit periods. When the parallel mode operation is used, all n digits of the numbers X and Y appe simultaneously in the same digit period. However, successi digits of increasing significance in the answer must still appe sequentially in time, because of the need to propagate the 'carr from one adder stage to the next stage of higher significance The delay between successive digits in the answer can, howev

The authors are in the Electrical Engineering Laboratories, Manchester University.

* Richards, R. K.: 'Arithmetic Operations in Digital Computers' (Van Nostra 1955).

be made small and less than a digit period, so that the parallel mode of operation still offers a speed advantage over the more economical serial type of addition. In practice, with computing machines having digit periods of 6–10 microsec it is relatively easy to ensure that all the answer digits, typically 40 in number, can be evaluated in one digit period. However, in computing machines now being designed, where the digit periods are of the order of 0·1 to 0·2 microsec, it becomes impossible to complete the addition in one digit period, and the effectiveness of the parallel mode in increasing the speed of addition is thus reduced. The following techniques indicate means for minimizing the addition time.

When two n-digit numbers are added together it is unlikely that a 'carry' propagation over n places will occur. At the Institute of Advanced Study, Princeton,* it has been estimated that, on the average, the maximum length of a '0's' or a '1's' 'carry' sequence in a 40-digit addition of random numbers is only 5·6 stages, Thus, if the adder could be designed to terminate the addition period at the end of the longest 'carry' sequence, an average saving in 'carry' propagation time of a factor of seven could be achieved.

The addition time during the course of a problem can also be effectively reduced by storing the 'carry' in a separate register and not permitting it to propagate at each addition. Assimilation of the partial sum with the 'carry' need occur only when the complete answer has to be standardized, sent to another storage location or tested for sign, although in this latter case complete assimilation is not always necessary. The scheme saves time when more than two numbers have to be added, because for any group of numbers there need only be one 'carry' propagation time. This scheme is of particular use during multiplication when numerous subproducts have to be added together.†

Even though it may be brief, the 'carry' propagation time in a single stage is very important, because its effect is cumulative. A small change in this time will alter the overall addition time appreciably, and therefore in an engineering design it is important to minimize the amount of equipment in the 'carry' path. There are two approaches to this problem, and these are identified by the manner in which the carry to stage k is provided. In the first method a logical gating system provides the 'carry' in terms of all the less-significant x and y digits. As the stages become more significant the complexity of this gating increases. This scheme is expensive, but has been used by the National Bureau of Standards‡ and they achieve a complete addition over 52 stages in 1 microsec. In the second method the 'carry' is provided from the previous stage, k-1, and that stage in its turn acquires its 'carry' from stage k-2, etc. The circuit to be described uses this method of operation.

A simple logical diagram of three stages of the adder utilizing on/off switches is shown in Fig. 1. These switches are closed when the logical operations appropriate to each switch occur. It should be noticed that the logical operations involve only the digits of x and y appropriate to that particular stage, and thus no switches have to be altered as a result of 'carry' propagation. In any one stage the switches T_1–T_3 are operated in a mutually exclusive manner so that normally there can be no interaction between the various sources which define the voltage level of the 'carry' path. If the switches are similar to relay contacts, the propagation time of a pulse through several switches in series corresponds merely to the time for a pulse to pass along a length of wire and is therefore for most practical purposes instantaneous. In this case the addition time is that appropriate to the setting time of the slowest switch.

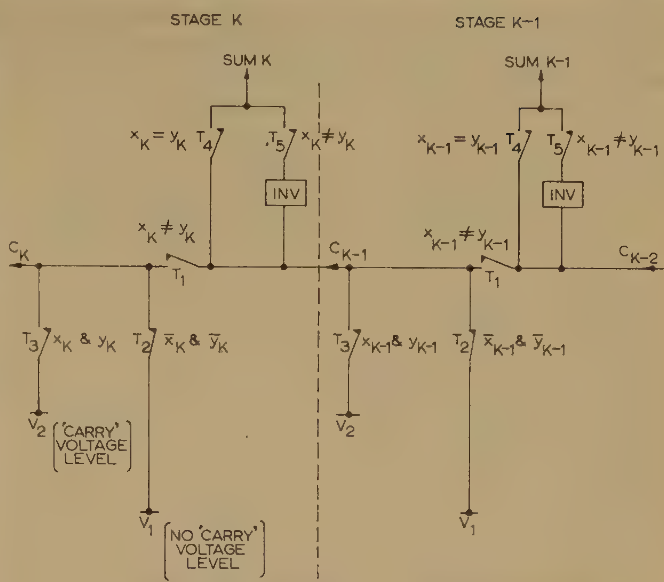


Fig. 1.—Simple logical diagram of two adder stages using on/off switches.

In a high-speed computer, relays are, of course, far too slow and they must be replaced by some form of electronic switch. These switches, however, suffer from several defects as compared with a mechanical contact: there is a significant voltage drop across the switch and also a current drain, so that the source of the 'carry' indication is subject to a load which varies in accordance with the number of switches operated. Both or either of these effects may limit the maximum number of electronic switches which can be connected in series. Diode 'and' gates have been used as switches, but the number which may be connected in series is limited to three or four by the rapid change in d.c. level of the transmitted pulse due to voltage drop across the diodes. A diode switch of this type allows information to pass only in one direction.

A junction transistor can be made to resemble closely the 'off' or 'on' positions of a mechanical switch by operating it in the cut-off or saturated condition. In the saturated condition the impedance between the emitter and the collector is very low, and the voltage drop is also small because it is essentially the difference in voltage drop across two conducting diodes. In this condition information can be passed in either direction through the transistor, provided that it does not significantly disturb the transistor's saturated condition. The surface-barrier transistor type 2N240 is satisfactory as a switch, but the micro-alloy diffused-base transistor type 2N501 is even better because of its increased current gain, current and voltage ratings and frequency response.

The simple experiment shown in Fig. 2 causes the square wave produced at the collector of T_0 to be transmitted through transistors T_1–T_{10}, all in series and operated in the saturated condition. The 2-volt movement of the emitter, collector and base potentials of each of these saturated transistors alters the base current only by approximately 10%, and this does not significantly alter the saturated conditions of the transistors. The waveforms from the collectors of T_1 and T_{10} occur simultaneously in time when viewed on an oscillograph range indicating 20 millimicrosec per centimetre.

A repetition of this simple experiment with 18 stages separated on individual plug-in boards results in a transmission delay of

* Gilchrist, B., Pomerne, J. H., and Wong, S. Y.: 'Fast Carry Logic for Digital Computers', *Transactions of the Institute of Radio Engineers*, 1955, EC-4, p. 133.

† 'On the Design of a Very-High-Speed Computer', Report No. 80, University of Illinois, Digital Computer Laboratory, October, 1957, p. 194.

‡ Weinberger, A., and Smith, J. L.: 'A One-Microsecond Adder using 1Mc/s Circuits', *Transactions of the Institute of Radio Engineers*, 1956, EC-5, p.65.

approximately 20 millimicrosec. It is thought that careful design of the boards to minimize the 5 ft length of wire in the 'carry' path will reduce this delay probably by a factor of two.

The practical circuit of two adder stages is shown in Fig. 3.

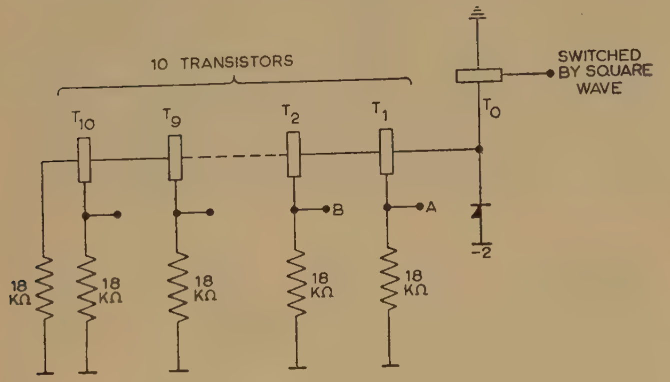


Fig. 2.—Experimental circuit.
All 18-kilohm loads are connected to a common supply at − 18 volts.

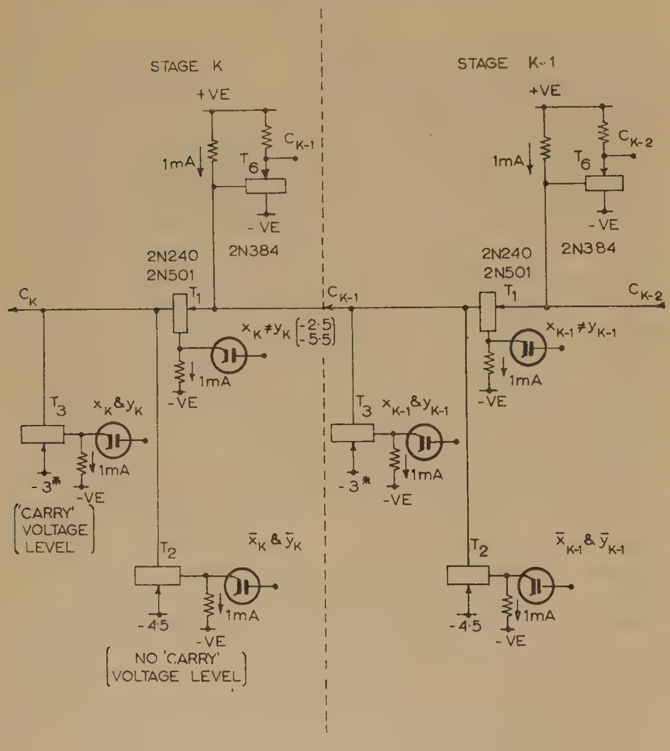


Fig. 3.—Practical circuit of two adder stages.

The 'sum' logic is not indicated, but it is driven from the emitter-follower T_6 which has been inserted to minimize loading on the 'carry' path. The logical operations which control the switching of transistors T_1–T_3 are carried out by simple diode logic using type OA47 diodes and both output phases of the flip flops storing x and y. The diode system operates extremely rapidly, and the gate outputs are emitter-followed prior to driving the bases of T_1–T_3. The amplitude of the function waveforms is from −2·5 to −5·5 volts. To ensure that the switch transistors remain under the control of the function waveforms, the 'carry' level must remain within these limits, and levels of −3 volts ('carry') and −4·5 volts (no 'carry') have been specified. Therefore a step 1·5 volts in amplitude is generated whenever a 'carry' occurs or disappears. When 'carry' propagates between stages, it does so via the T_1 transistors, a large number of which therefore appear in series.

The 1 mA positive leak connected to the emitter of T_1 transistors provides sufficient current for these to saturate without the need for any current to be provided along the 'carry' path. The transistor defining the 'carry' indication passes a large current only transiently, in order to charge the capacitances along the 'carry' path. The T_1 transistors bottom to a very low voltage, because their direct current is negligible; typically, the drop across ten T_1 transistors of either type is approximately 0·3 volt. In practice, after ten stages an emitter-follower is used to reconstitute the voltage level by using the voltage difference across the base-emitter junction, which will vary by a maximum of 30%. If the emitter-follower had raised the 'carry' level by exactly the correct amount, i.e. 0·3 volt, then in the next ten T_1 stages there is only the same 0·3-volt fall in voltage level, so that no cumulative loss occurs in subsequent groups of ten. It is hoped that the 'carry' in this way will be able to propagate along 40 stages without the need for reamplification.

Ideally, T_1–T_3 operate in a mutually exclusive manner, but under fault and transient conditions this need not apply. For example, if switches T_2 and T_3 were both closed, the −4·5- and the −3-volt lines would be connected together. To avoid excessive current the −3-volt line is limited in the amount of current it can supply.

When a 'carry' propagates along the 18 stages which have been constructed, the delay plus rise time to the eighteenth stage is approximately 80 millimicrosec, comprising the switching time of transistors, including any delay due to generation of the switching waveforms, and the transmission delay of approximately 20 millimicrosec.

The latter is easily seen from the 'carry' indication at the second and eighteenth stages. When the number of stages is increased the transistor switching time will remain constant, but the transmission time will be increased. If the former time is 60 millimicrosec and the latter approximately 1 millimicrosec per stage, it would appear feasible to carry out a complete 40-stage parallel addition in 0·2 microsec.

TRANSISTOR LOGIC USING CURRENT SWITCHING AND ROUTING TECHNIQUES AND ITS APPLICATION TO A FAST 'CARRY' PROPAGATION ADDER

By L. P. Morgan, B.A., and D. B. Jarvis, B.Sc.

The fastest method of operating transistors is in non-saturating circuits using a voltage drive on the base or a current drive into the emitter. A set of logical circuits using these principles has been suggested by Yourke.* This system essentially uses a long-tailed pair, the conducting transistor being selected by a voltage drive on the base. Interconnection between elements is by passing the output current from the collector of a unit through a low resistance in the base of one transistor of the next unit. The system is d.c. connected, and the problem of the increasing supply voltage as more stages are cascaded can be overcome by using *n-p-n* transistor or Zener diodes.

When two or more stages are cascaded in this system, significant delay is caused by the fact that the output from one transistor must overcome the hold-off voltage of the following stage. In several instances, however, this delay can be reduced if the output current is allowed to pass directly into the emitter of the following stage. With this modification a parallel adder with a fast 'carry' propagation has been developed.

The circuit of a single stage of the adder is shown in Fig. 1,

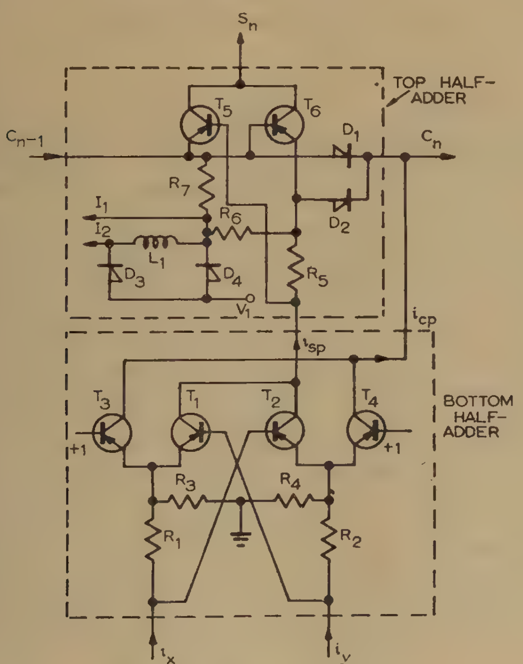


Fig. 1.—Single stage of adder.

the inputs i_x and i_y being currents from a high-impedance source. The stage may be considered as two half-adders. The operation is as follows:

For the bottom half-adder,

(*a*) If both input currents are zero there are no output currents.

(*b*) If either one or the other input is '1', either T_1 or T_2 conducts and there is a partial sum current, i_{sp}, from the bottom half-adder.

(*c*) If both inputs are '1's', T_1 and T_2 are cut-off and the currents flow through T_3 and T_4 to give a 'carry' output, i_{cp}.

For the top half-adder,

(*d*) If there is no 'carry' current from the preceding stage and i_{sp} is '0', there is no output.

(*e*) If there is either a 'carry' or a partial sum current, either T_5 or T_6 conducts. The input currents are prevented from flowing through D_1 and D_2 by the supply potential V_1 being approximately 0·5 volt more positive for each successive stage.

(*f*) If both inputs are '1's', T_5 and T_6 are cut off and the 'carry' and sum are propagated through D_1 and D_2.

The voltage developed on the 'carry' line at any given stage can be determined approximately by adding 1 volt for each stage through which the 'carry' is propagated beyond the given stage. This has two disadvantages: first, not more than six or seven stages can be cascaded without a level-changing unit; and secondly, the resistors R_6 and R_7 (which must be of low value, since they are in series with the bases of T_5 and T_6) shunt the 'carry' line and attenuate the propagated 'carry' current. The latter disadvantage can be overcome as follows:

When a carry is to be propagated, the input current to the top half-adder is of the order of $2i$, and the propagated 'carry' current need only be i; therefore, if the current through R_7 and R_6 is limited to i, sufficient output current will always be obtained. This limiting is achieved by forward biasing D_3 and D_4 by I_1 and I_2 (when $I_1 + I_2 = i$). The inductance L_1 speeds the 'carry' propagation by restricting the current to I_1 during the transient, where I_1 is the peak base current.

With this circuit a 'carry' propagation time over six stages of approximately 150 millimicrosec is achieved with OC44 transistors, and of about 40 millimicrosec with OC170 transistors. The delay through a level-changing unit using OC44's is approximately 70 millimicrosec, so that using blocks of six stages, a 52-stage adder would have a 'carry' propagation time of approximately 2 microsec.

This time can be reduced if a 'carry' skip technique is used over each of the groups of six stages. The logical block diagram for this skip technique is shown in Fig. 2. Fig. 3 shows the modified

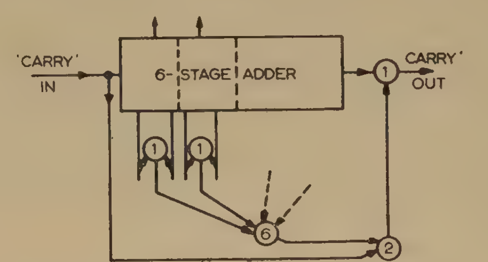


Fig. 2.—Logical circuit for skip technique.

bottom half-adder; this provides the diode 'or' circuit and one transistor of the 6-input 'and' circuit.

The level-changing and skip unit is shown in Fig. 4. The propagated 'carry' current passes through T_{13} and a corresponding change of current occurs in the resistor connected to the base of T_8. This switches the constant current I_5 through T_9 and thence to the carry input of the next six stages. The skip facility is provided by T_{10}–T_{12}. If a 'carry' is to skip the following six stages, all the transistors connected to the line D, except T_{12}, will be cut off, so that the constant current I_4 will pass through T_{12}. If there is a 'carry' from the preceding six stages, this current will be routed through T_{10}, and thence to the point F of the next level-changing unit. The currents I_6 and I_5 are chosen so that up to two units of current can be fed

* Yourke, H. S.: 'Millimicrosecond Transistor Current Switching Circuits', *Transactions of the Institute of Radio Engineers*, 1957, CT4, p. 236.

Mr. Morgan and Mr. Jarvis are at the Mullard Research Laboratories.

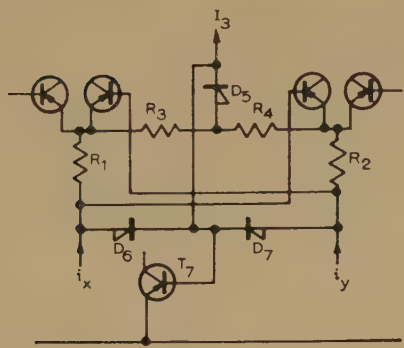


Fig. 3.—Modified bottom half-adder.

into the transistor T_{13}, the only effect being that the base of T_8 is taken more positive.

The skip unit introduces a delay of approximately 50 millimicrosec, so that it is now possible to propagate the 'carry' in less than 1 microsec. However, this is only the propagation time for a '1', and the time taken to propagate a '0' after a '1' has been propagated is still 2 microsec. Several ways of overcoming this have been suggested, one of which is to turn off the current I_5 whenever the input registers are changed.

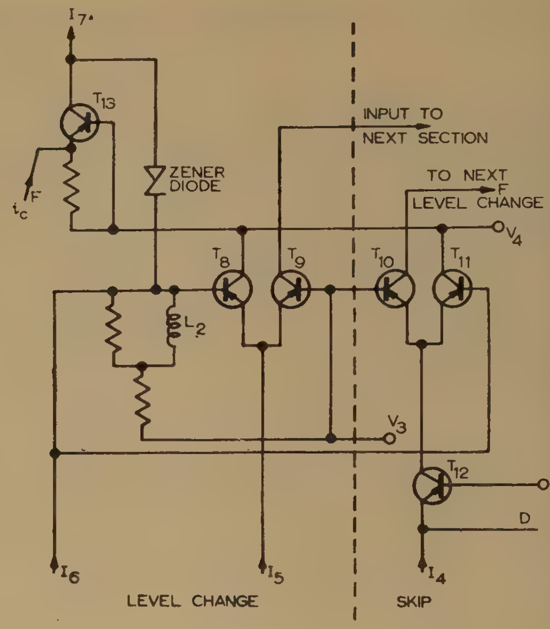


Fig. 4.—Level-changing and skip unit.

THE DESIGN PRINCIPLES OF THE NEURON AND RESONANT-CIRCUIT LOGICAL ELEMENTS

By G. G. Scarrott, Associate Member, K. C. Johnson, M.A., G. Haley, B.Sc., and R. Naylor, M.Sc., Graduat

Several different designs of logical element have already been proposed for use as 'building bricks' in the construction of digital-computing equipment. The two systems to be described are based on the use of transistors for the power gain, and both make use of wound ferrite cores as linear transformers rather than as square-loop storage devices. These transformers give impedance matching and d.c. isolation, but are principally employed to allow the use of 'ballot-box' logical gating in place of the conventional Boolean system. This allows a single gate circuit to be used to perform any of the fundamental logical operations, and this flexibility enables each gate to be permanently associated with a power-amplification and pulse-restandardization stage. Thus only a single basic element is, in principle, necessary to build up the logical parts of a computer, although, in practice, some variations may be introduced to improve the overall economy of the system.

The pulses transmitted from one element to the next are of current, rather than voltage, in both systems and are synchronized to a standard clock frequency of 500 kc/s. The logical properties are obtained by summing these current pulses on a transformer at the input of each element by the use of a number of separate primary windings, each of which may be connected to carry an incoming current in either a positive or negative direction. The restandardization circuit is arranged to generate an output pulse of standard current amplitude and timing if, and only if, the total effective current at the primaries is positive; if the total primary current is zero or negative, no output pulse is generated. The circuits are so designed that the actual number of pulses contributing to the total current does not matter, although it will not normally exceed about three standard units in either direction. The likeness of this process to voting in an election explains the name 'ballot-box' logic.

The authors are with Ferranti, Ltd

To obtain an 'or' gate with this system, all the various inpu are connected positively to the input transformer and a curre at any one suffices to ensure an output pulse. For an 'an gate the various inputs are again connected positively, but, addition, a constant stream of pulses is made to flow negatively connected windings, such that only coincident puls at every input are able to make the necessary positive total generate an output pulse. To make an inverter the input puls are connected to flow negatively in one winding while a consta stream of unit strength is connected positively to another, that output pulses are obtained only when the input has pulse. Thus, this type of gate can perform all these bas logical operations merely by suitable connection of the inpu to a single identical circuit; moreover, it will be seen that t examples given do not by any means exhaust the possibilities.

Up to this point the neuron and resonant-circuit systems ha been considered together, but they differ considerably in t details of the restandardization circuit. Hence it will be co venient to divide and discuss the two systems separately fro here onward to avoid confusion.

In the basic neuron circuit the standard current pulses are 9 mA amplitude and last for one-half of the standard digit tim The digit period is thus divided into two equal portions, ea 1 microsec in length, which are known as the charging peri and the output period and are defined throughout the equipme by means of timing waveforms. These are generated in tv phases, A and B, such that the charging period of one phase simultaneous with the output period of the other and vi versa.

In the operation of the neuron circuit the current from secondary winding on the input transformer flows into arrangement of three diodes. These are connected so that, the total input current is positive, in the logical sense alrea described, this current will charge a capacitor during

charging period from one defined voltage level to another. If the total current is zero or negative, the capacitor is left at the former voltage and the third diode absorbs the current and prevents the generation of any undue voltage across the windings. During the following output period the timing waveforms are arranged so that the output transistor is switched on to carry the standard current of 9 mA if, and only if, the state of charge of the capacitor has been altered by a positive total input; at the same time the transformer is allowed to make a voltage swing so that it is d.c. restored by the beginning of the next charging period. A very useful feature of this arrangement is that it will accept input currents during either the charging period or the output period, giving an overall delay of either a half or a whole digit time respectively. Thus the one standard design of unit is able to perform powerful logic with only the half-digit delay or alternatively to delay one whole digit time and still give some useful logical facilities.

This basic circuit is varied to make the double-entry neuron, which has two input transformers but gives an output if either has a positive total input. Also there is a booster circuit, which enables the standard currents to be distributed more widely when required, and both electromagnetic and magnetostriction delay devices have been integrated into the system. A computer has been made using these logical elements and has been working successfully.

The logical properties of the resonant-circuit device are provided by the summation transformer as described above. One of the secondary windings of this transformer is formed into a damped resonant circuit with a Q-factor of unity. If the total effective current at the primaries is positive, the voltage across the tuned circuit follows the usual response of a resonant circuit to an input current pulse, and a clock waveform added at the time of the voltage overswing causes a standardized current pulse to be issued by a long-tailed pair formed by a transistor and diode. In this case, the standardized current pulses have an amplitude of 18 mA and duration of one-quarter of the digit period. In contrast to the neuron logical elements, the resonant-circuit logical elements use only one transistor per stage, require a simple clock waveform and three units can be mounted on a printed circuit which will accommodate only two neurons.

Since there is only one information storage mechanism in a resonant-circuit element, it is impossible to use this element to provide a whole digit delay. The resonant-circuit elements have the added disadvantage that their transformers are of a more complex design.

DISCUSSION ON SPECIAL ASPECTS OF LOGICAL DESIGN—II

The discussion during this session was almost entirely concerned with the contribution by Messrs. Scarrott, Johnson, Haley and Naylor. **Mr. L. J. Bental** felt that the logical element described was more complicated than it need be. He was familiar with an element with similar properties which made use of the square-loop property of a ferrite material and used only about a quarter the number of components. **Mr. G. G. Scarrott** replied that, because his element made use of the linear characteristic, the power consumption was far less than that required to switch the ferrite and the speed of operation was high—500 kc/s in the present design (replying to a further question Mr. Scarrott thought that the upper limit determined by transit times within the transistor and winding capacitances would be about 2 Mc/s). Several questions asked by **Mr. D. Eldridge** produced the information that some eight elements could be driven from one output, and that when driving along long wires it might be necessary to retime pulses before any further logical operations are performed. Some discussion took place on whether a voltage-pulse or current-pulse technique was better, and it was felt that prejudice was often the determining factor in this. The logical complexity of the element was defended by Mr. Eldridge, who felt that in a computer design it was usually possible to use the element efficiently, so that an overall saving of time and equipment resulted.

The only other contribution of any substance during the discussion period was given by **Mr. G. B. B. Chaplin**, who delivered a short account, well illustrated by slides, of some work that he had recently done in the design of a fast parallel multiplier. His design involved logical elements involving analogue addition, voltage addition being chosen because it was thought to be faster than the alternative of current addition.

In his closing remarks of this second session the Chairman, **Mr. E. P. G. Wright**, commented on the nomenclature of computers. When at a previous session mechanical devices were described they were all 'high-speed' devices—high-speed printers, high-speed readers and so on. The present session had however dealt with 'fast' equipments—fast address, fast logic. He wondered what reaction one would have from a railway official if one asked when the next 'high-speed' train would leave.

621.317.34:621.315.2

The Institution of Electrical Engin
Paper No. 30
Sept. 1

A METHOD OF LOCATING DRY JOINTS IN TELECOMMUNICATION CABLES

By J. RHODES, M.B.E., B.Sc.(Eng.), Associate Member.

(*The paper was first received* 20*th February, and in revised form* 24*th April,* 1959.)

SUMMARY

A description is given of a method of locating dry joints in telecommunication cables. The method is based upon the non-linear properties of dry connections and has been used successfully on both land and submarine cables.

(1) INTRODUCTION

'Dry connections' might be defined as soldered connections in which the solder whilst molten has failed to wet the surfaces of the connection, or alternatively connections which have been subject to a tension sufficient to break the solder film.

They occasionally occur in telecommunication cables and are liable to cause intermittent disconnections, noise, variations in attenuation, and intermodulation. These faults can be particularly troublesome on modern coaxial cables carrying wideband transmission systems of telephony or television. Methods which have been used to locate this type of fault range from the use of cable pulse-testing equipment, with which an attempt is made to observe the impedance irregularity produced by the fault, to vibrating all the joints in the cable in turn whilst, at the same time, observing the d.c. resistance from the end of the cable. The first method is not very sensitive and if a high pulse power is used the fault is often temporarily sealed. The second is an expensive and time-consuming method, particularly if the joints are buried.

(2) DESCRIPTION OF METHOD

The method is based on the non-linear property of a dry connection and the localization is obtained by comparing the phase of the harmonic produced at the fault to that of a harmonic generated at the terminal. It was developed when a submarine cable became slightly non-linear and produced harmonics in the transmission band which, although of very low amplitude (approximately 130 dB below the fundamental) were nevertheless sufficient to make the transmission system unworkable owing to the high level differences between different directions of transmission. Its use has now been extended to underground coaxial cables. Fig. 1 shows the principle of the apparatus used. F_1 and F_2 is a high-pass/low-pass filter combination having air-cored inductances to avoid any harmonic production in the coils. At the input to the low-pass filter (transmit side) a variable-frequency oscillator is connected and at the output of the high-pass filter (receive side) a suitable wideband amplifier (or amplifiers) to give the required gain is used, followed by a measuring set. This measuring set should be of a narrow-band selective type to obtain maximum sensitivity against thermal noise. The cable to be tested is connected to the junction of the filter combination. Also at this point an adjustable non-linear element may be connected as required during the test.

The oscillator is first set to a frequency f just above $\frac{1}{2}f_c$ (f_c being the cut-off frequency of the high-pass filter) and the gain on the receive side is adjusted until the second harmonic, $2f$, generated at the fault, produces a suitable reading on the selective

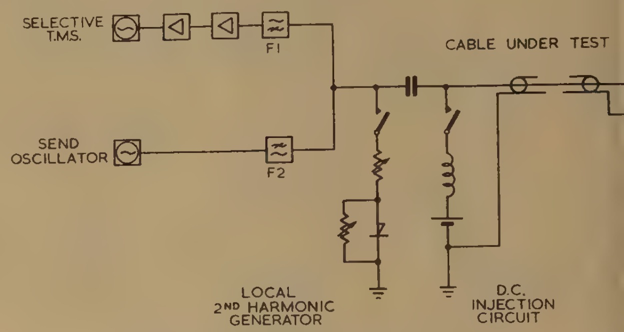


Fig. 1.—Principle of apparatus used.

measuring set. The cable is then disconnected, and the junct of the two filters is terminated by a resistance equal to the ca impedance (a wire-wound resistor should be used to av harmonic production) together with the adjustable non-lir element. A copper-oxide or germanium rectifier as used modulators is suitable for this purpose. This non-linear elem is then adjusted by means of a series and shunt resistance u the second harmonic produced by it is of the same magnit (as measured on the selective measuring set) as that produ by the cable. The terminating resistance is then removed f the common point of the directional filters and the cabl reconnected while still keeping the non-linear element conne in parallel at this point. A series of readings is then taken increasing frequencies up to the cut off of the low-pass filter

(3) THEORY OF METHOD

The readings of the selective measuring set are the vector of the second-harmonic signal generated locally by the non-li element and the second-harmonic signal generated at the fa As the test frequency is increased the phase of this latter si will change whilst the phase of the locally generated seco harmonic signal remains at zero. Consequently, as the signals move in and out of phase, a curve consisting of a se of peaks similar to that shown in Fig. 2 will be produced. Δf is the mean frequency between consecutive peaks and v_l v_h are the mean phase velocities in the low- and high-frequ bands respectively,

$$d = \frac{v_l v_h}{2\Delta f(v_l + v_h)} \text{ (see Section 9)}$$

The method is similar in some respects to the well-kn impedance/frequency method of locating an impedance irr larity. However, there is an important difference since, by b able to adjust the level of the second harmonic which is gener locally at the terminal to the same level as that returned f the fault, the sensitivity is considerably increased. Th because the amplitudes of the two signals whose phases compared are equal, and the only limit is the ability of measuring equipment to recognize the peaks of the cu against the background of resistance noise. In the no

Written contributions on papers published without being read at meetings are invited for consideration with a view to publication.

Mr. Rhodes is in the Post Office Engineering Department.

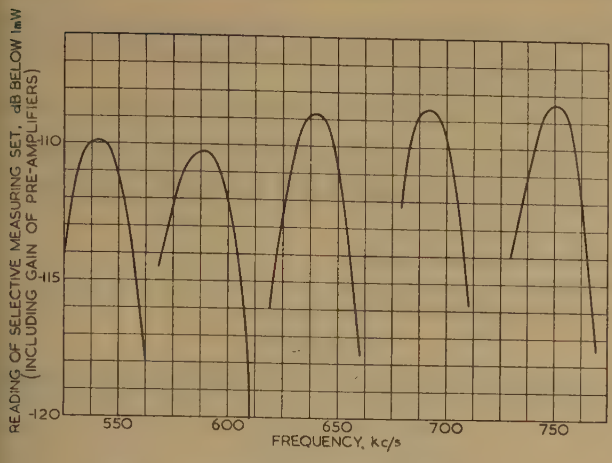


Fig. 2.—Typical curve of peaks obtained.

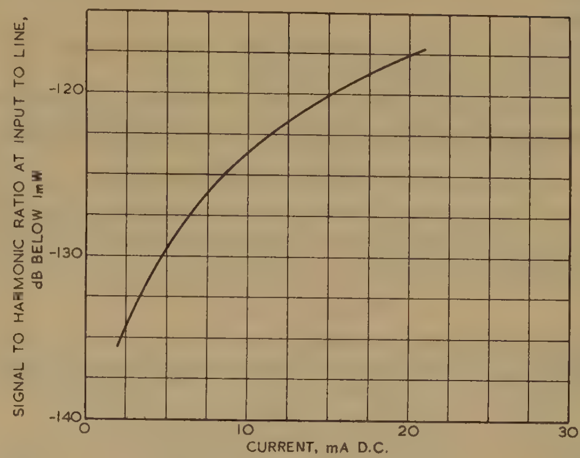


Fig. 3.—Increase in second harmonic as direct current is applied.

impedance/frequency test the phase of a signal of low level reflected from the irregularity is compared with the phase of the total signal sent to the line, which consequently greatly limits the sensitivity.

(4) CHOICE OF FREQUENCY RANGE

For normal purposes it is sufficient to choose the frequency of the sent signal, so that a curve with about three peaks is obtained from which the average frequency between successive peaks may be calculated. Furthermore, the frequency range of the fundamental must be less than one octave. The frequency between peaks is, of course, dependent upon the distance to the fault, and therefore some arbitrary value must be chosen. For a standard 0·375 in air-spaced coaxial cable laid in repeater sections of six miles a suitable range would be such that a 3-peak curve is obtained from faults one mile or more from the sending end. Faults nearer than this may be localized by testing from the distant end. This avoids having a range of special directional filters, which are the only specialized items of plant required for the test. For the case quoted, the cut-off frequency of the low-pass filter should be about 340 kc/s, although a higher figure can be used with very little loss of sensitivity. This allows for a 20% gap between the cut-off frequencies of the low-pass and high-pass filters.

(5) ACCURACY OF LOCATION

The accuracy of location is dependent upon the accuracy of the oscillator, the phase shift which may occur at the fault and the constancy of the velocity over the frequency band used. In practice, joints in underground cables in the United Kingdom are generally at intervals of 100–176 yd, and no difficulty has been experienced in locating a faulty joint in a 6-mile repeater section. It is estimated that, if great care is taken and specially calibrated apparatus is used, location to within a few yards would be possible.

(6) USE OF DIRECT CURRENT TO INCREASE SENSITIVITY

The sensitivity of the method is appreciably increased if the harmonic production at the fault is increased. Since the harmonic is caused by the rectifying action at the point of contact, the surfaces of which will have a very thin oxide coating, the generation of the second harmonic can be considerably increased by biasing the contact. This may be done by passing a direct current through the cable, as shown in Fig. 1. Fig. 3 shows the increase in second harmonic which was obtained from a particular fault when various values of direct current were passed. The fault in this case was found to be a badly soldered sleeve on the outer conductor in a 0·375 in underground coaxial cable. Where the harmonic level has been found to vary rapidly during testing, the application of a d.c. bias considerably stabilizes the signals.

(7) LIMIT OF SENSITIVITY

The limit of sensitivity depends upon the power sent to line, the distance to the fault, the second-harmonic generation at the fault and the bandwidth of the measuring set used, assuming that it has ample gain. This is not a very onerous requirement as it may be preceded by a suitable amplifier. It is usually possible to transmit the fundamental signal at +20 dBm, i.e. reference level of 1 mW, and with a selective measuring set of 200 c/s bandwidth to measure second-harmonic signals at a level of −140 dBm from the line. Thus second-harmonic levels of 160 dB below the outgoing fundamental signal are capable of being measured. For a 6-mile repeater section of 0·375 in coaxial pair cable the range of the fundamental signal required is 205–340 kc/s. At 340 kc/s the attenuation of the fundamental to, and of the second harmonic from, a fault at the far end of the section is some 30 dB. Therefore, a fault at the far end producing a second harmonic of 130 dB below the fundamental can be located. Faults with a higher harmonic margin may be located if they are nearer the testing end up to values of the order of 155 dB at a distance of one mile.

(8) CONCLUSION

The author wishes to thank his colleagues in the Post Office Engineering Department who assisted in the development of the method and to the Engineer-in-Chief of the Post Office for permission to publish the paper.

(9) APPENDIX

Derivation of the expression for location of the fault

List of Symbols.

f_n, f_{n+1} = Frequencies at which successive peaks occur.
ω_n, ω_{n+1}, etc. = $2\pi f_n$, $2\pi f_{n+1}$, etc.
Δf = Mean frequency interval between successive peaks.
v_{fn} = Phase velocity over the cable at frequency f_n.
v_l = Mean phase velocity over the low-frequency band.
v_h = Mean phase velocity over the high-frequency band.
β_{fn} = Phase-change coefficient at frequency f_n.
d = Distance to the fault.

Proof.

The proof is concerned only with relative phase. Let us consider first a signal at frequency f_1.

If $\sin \omega_1 t$ is the fundamental signal at frequency f_1 applied to the cable, then at the point of the fault it will be represented by

$$\sin (\omega_1 t - \beta_{f1} d)$$

If the fault has a non-linear characteristic of the form $v = ai + bi^2$ the second-order harmonic generated at this point will be $\sin^2 (\omega_1 t - \beta_{f1} d)$

i.e. $$\cos (2\omega_1 t - 2\beta_{f1} d)$$

if the d.c. component is omitted.

This signal on its return to the sending end will be represented by

$$\cos [2\omega_1 t - d(2\beta_{f1} + \beta_{2f1})]$$

Now $$\beta_{f1} = \frac{\omega_1}{v_{f1}} \text{ and } \beta_{2f1} = \frac{2\omega_1}{v_{2f1}}$$

and the expression becomes

$$\cos \left[2\omega_1 t - d\left(\frac{2\omega_1}{v_{f1}} + \frac{2\omega_1}{v_{2f1}}\right)\right]$$

This signal combines with the second harmonic $\cos 2\omega_1 t$ ge rated locally and will produce the characteristic peaks as frequency is increased.

Thus for the next peak at frequency f_2 the signal is rep sented by

$$\cos \left[2\omega_2 t - d\left(\frac{2\omega_2}{v_{f2}} + \frac{2\omega_2}{v_{2f2}}\right)\right]$$

and for two adjacent peaks

$$2d\left(\frac{\omega_2}{v_{f2}} + \frac{\omega_2}{v_{2f2}}\right) - 2d\left(\frac{\omega_1}{v_{f1}} + \frac{\omega_1}{v_{2f1}}\right) = 2\pi$$

Substituting $\Delta f = f_2 - f_1$, v_l for v_{f1} and v_{f2}, and v_h for v and v_{2f2}:

$$d = \frac{v_l v_h}{2\Delta f(v_l + v_h)}$$

If the frequency used is sufficiently high for the phase veloc to be assumed constant,

$$d = \frac{v}{4\Delta f}$$

621.385.1.029.6

The Institution of Electrical Engineers
Paper No. 2789 E
Jan. 1959
©

A SIMPLE INVESTIGATION OF THE CROSS-MODULATION DISTORTION ARISING FROM THE PULLING EFFECT IN A FREQUENCY-MODULATED KLYSTRON

By D. T. GJESSING, B.Sc.

(*The paper was first received* 14*th July*, 1958, *and in revised form*, 6*th October*, 1958. *It was published in January*, 1959, *and was read before the* Electronics and Communications Section 23*rd February*, 1959.)

SUMMARY

The paper gives a simple theoretical approach to the problem of cross-modulation distortion arising from frequency pulling in, or in direct conjunction with, a frequency-modulated klystron. An expression for the distortion spectrum relative to the fundamental component is derived, and on the basis of this expression a family of curves illustrating the severity of the distortion is computed and plotted on a logarithmic scale. It is shown that even small mismatches on the far end of a relatively short feeder into which the klystron is working give rise to serious crosstalk which, if not remedied, sets a limit to the quality of the modulator unit.

LIST OF PRINCIPAL SYMBOLS

λ = Free-space wavelength, m.
λ_g = Guide wavelength, m.
ρ = Reflection coefficient of load.
θ = Phase angle of load, rad.
l = Length of feeder connecting the modulator to the load, m.
a = Largest cross-sectional dimension of guide, m.
ϕ = Phase of load referred to the modulator, rad.
c = Velocity of light, m/s.
ω_a = Modulator frequency, rad/s.
f_c = Carrier frequency, c/s.
f_d = Frequency deviation, c/s.
ω_d = Frequency deviation, rad/s.
f_p = Pulling frequency, c/s.
J_n = nth-order Bessel operator.
D_n = nth-order harmonic distortion component relative to the fundamental.
p = Pulling figure, c/s.
S = Voltage standing-wave ratio of the load.
S_s = Distance of the sink region from the operating point measured in v.s.w.r. units.
ρ_s = Distance of the sink region from the operating point measured in units of ρ.
$q = \left(1 - \dfrac{\rho}{\rho_s}\right) \Big/ \left(1 + \dfrac{\rho}{\rho_s}\right)$.
T_g = Twice the group delay of the line.

(1) INTRODUCTION

During the last few years the problem of noise sources in multi-channel f.m. systems has been treated both analytically and experimentally in considerable detail by several authors. The phase distortion (echo distortion) caused by multi-reflections in long feeders and the problem of amplitude distortion introduced by a non-linear modulator characteristic have been investigated. The third source of interference in the modulator, namely that due to the pulling effect in a modulated klystron, does not, however, seem to have been considered very seriously.* Admittedly this effect is negligible from the point of view of crosstalk when the klystron is being used purely as a modulator working directly into a power amplifier or coupled to a load isolator of some sort, e.g. a ferrite non-reciprocal network, but for many applications one wishes to dispense with additional equipment and requires the modulated klystron to work as a transmitter connected directly to the aerial system through a feeder. In practice the effective length of this feeder may be fairly large, and even small mismatches on the aerial side of the feeder will give rise to large-scale non-linearities which, if not remedied, limit the quality of the modulator unit.

* A paper dealing with noise in a klystron modulator has since been published: Rice, S. O., and Curtis, H. E.: 'Interchannel Interference due to Klystron Pulling', *Bell System Technical Journal*, 1957, 36, p. 645.

Mr. Gjessing is at the Norwegian Defence Research Establishment, Bergen.

(2) SIMPLE BASIC THEORETICAL CONSIDERATIONS

It is well known that if the load presented to a klystron oscillator is varied, either in phase or in magnitude, there is a corresponding change in the oscillation frequency of the valve. A Riecke diagram is a convenient representation of the stability of the klystron for varying loads. With this diagram as a basis, a simple analysis of the degree of crosstalk introduced in a modulated klystron as a result of frequency-pulling effects will be presented.

Consider a klystron connected to a mismatched load through a long feeder. Let the normal oscillation wavelength of the valve be λ, the length of the feeder connecting the valve to the load l, and the reflection coefficient of the load referred to the plane of the load ρ.

For the reflection coefficient referred to the valve,

$$\rho_1 = |\rho| \exp j\left(\frac{4\pi l}{\lambda_g} + \theta\right)$$

where

$$\lambda_g = \frac{\lambda}{\sqrt{\left[1 - \left(\frac{\lambda}{2a}\right)^2\right]}} \quad . \quad . \quad . \quad . \quad . \quad (1)$$

We assume that the reflection coefficient of the load, both in magnitude and in phase, is independent of the frequency within a limited interval and write for the phase of the load referred to the modulator,

$$\phi = \frac{4\pi l}{\lambda_g} + \theta \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

To evaluate the frequency pulling of the klystron caused by a variation in the angle of the load presented to it, we require the differential of the phase angle with respect to the frequency.

Now

$$\frac{d\phi}{df} = \frac{d\phi}{d\lambda_g}\frac{d\lambda_g}{d\lambda}\frac{d\lambda}{df} \quad . \quad . \quad . \quad . \quad . \quad (3)$$

Evaluating the appropriate differentials,

$$\frac{d\phi}{df} = -\frac{4\pi l}{c\sqrt{\left[1 - \left(\frac{\lambda}{2a}\right)^2\right]}} \text{ radians per cycle} \quad . \quad (4)$$

Consider now the klystron being frequency modulated by a signal whose amplitude is given by

$$V = V_m \cos \omega_a t$$

Employing the usual conception of a frequency-modulated signal, the instantaneous frequency is

$$f = f_c + f_d \cos \omega_a t$$

The variation in the phase of the load presented to the klystron modulator as a result of this frequency modulation is obtained from eqn. (4) and given by

$$\Delta\phi = \frac{4\pi l}{c\sqrt{\left[1 - \left(\frac{\lambda}{2a}\right)^2\right]}} f_d \cos \omega_a t \quad . \quad . \quad . \quad (5)$$

The appropriate relationships between the frequency pulling f_p and the fluctuations in the phase angle of the load can be determined by a more practical consideration employing the Riecke diagram of the valve. An expression must now be found giving the magnitude of the nth-order component in the frequency spectrum of the distorted signal relative to that of the fundamental component.

It will be seen from Fig. 1 that within a limited interval of

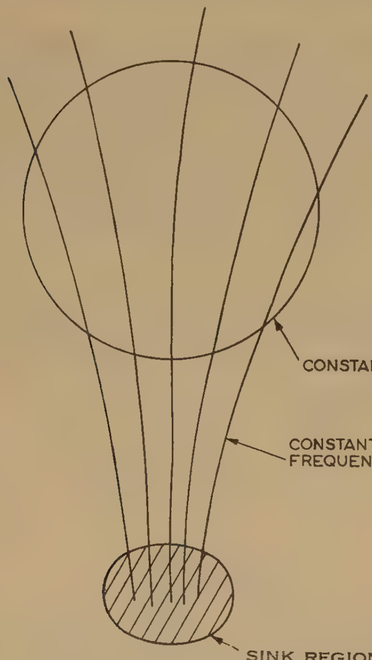


Fig. 1.—A typical Riecke diagram.

frequency and reflection coefficient the lines representing constant frequency can justifiably be regarded as a set of straight lines, converging somewhere in the sink region of the diagram. It seems justifiable to assume that the static and dynamic Riecke diagrams are identical and that the time required for the modulated wave to travel from the klystron to the mismatched load and back again is short in comparison with the period of the signal that is modulating the klystron. Referring to the idealized Riecke diagram in Fig. 2, where β is the maximum frequency-pulling that can be obtained for a given load mismatch, and the remaining symbols are as indicated in Fig. 3, the frequency pulling is given by

$$f_p = (\rho + \rho_s)A \tan \xi \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

where A is a constant of dimensions (frequency)/(reflection coefficient).

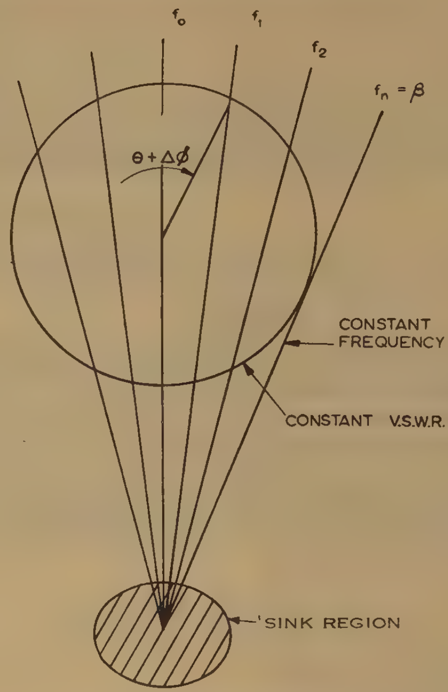


Fig. 2.—Idealized Riecke diagram.

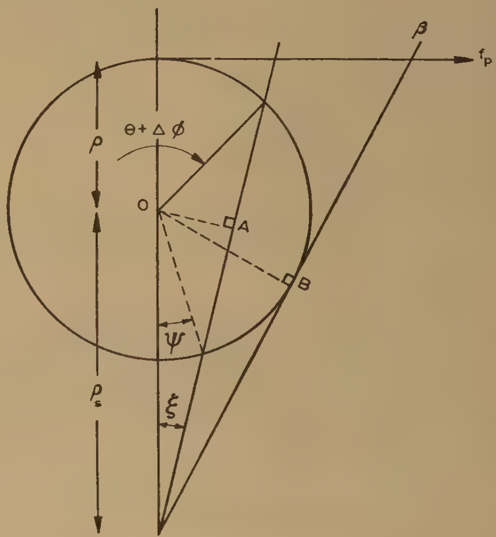


Fig. 3.—The geometry of the Riecke diagram.

When $\quad \tan \xi = \dfrac{\rho}{\sqrt{(\rho_s^2 - \rho^2)}}$

we have $\quad f_p = \beta = (\rho + \rho_s)A \dfrac{\rho}{\sqrt{(\rho_s^2 - \rho^2)}}$

i.e. $\quad A = \dfrac{\sqrt{(\rho_s^2 - \rho^2)}}{(\rho + \rho_s)\rho}\beta$

Substituting for A in eqn. (6),

$$f_p = \frac{\sqrt{(\rho_s^2 - \rho^2)}}{\rho}\beta \tan \xi \quad . \quad . \quad . \quad .$$

From Fig. 3, $\sin \xi = \frac{\rho}{\rho_s} \sin \left[\frac{(\Delta\phi + \theta) + \psi}{2}\right]$

and $$\xi = \frac{(\Delta\phi + \theta) - \psi}{2}$$

which by substitution yields

$$\tan \xi = \frac{(1 - q) \sin (\Delta\phi + \theta)}{(1 - q) \cos (\Delta\phi + \theta) + (1 + q)} \quad (8)$$

where $$q = \frac{1 - \rho/\rho_s}{1 + \rho/\rho_s}$$

By simple substitution for $\tan \xi$ in eqn. (7),

$$f_p = \frac{2\beta\sqrt{q} \sin (\Delta\phi + \theta)}{(1 - q) \cos (\Delta\phi + \theta) + (1 + q)} \quad (9)$$

Inserting the expression for $\Delta\phi$ from eqn. (5),

$$f_p = \frac{2\beta\sqrt{q} \sin (K \cos \omega_a t + \theta)}{(1 - q) \cos (K \cos \omega_a t + \theta) + (1 + q)} \quad (10)$$

where $$K = \frac{4\pi l}{c\sqrt{[1 - (\lambda/2a)^2]}} f_d = T_g \omega_d$$

Expanding (10) by the binomial theorem, neglecting powers higher than the second and rearranging,

$$f_p = \frac{2\beta\sqrt{q}}{1+q}\left\{\left[1 + \frac{1}{4}\left(\frac{1-q}{1+q}\right)^2\right] \sin (K \cos \omega_a t + \theta) - \frac{1}{2}\left(\frac{1-q}{1+q}\right) \sin (2K \cos \omega_a t + 2\theta) + \frac{1}{4}\left(\frac{1-q}{1+q}\right)^2 \sin (3K \cos \omega_a t + 3\theta)\right\} \quad (11)$$

Expanding $\sin (K \cos \omega_a t)$ and $\cos (K \cos \omega_a t)$ in Bessel functions,

$$\sin (K \cos \omega_a t + \theta) = \mathscr{I}\left[J_0(K) + 2\sum_{n=1}^{\infty} j^n J_n(K) \cos n\omega_a t \cos \theta\right] + \mathscr{R}\left[J_0(K) + 2\sum_{n=1}^{\infty} j^n J_n(K) \cos n\omega_a t \sin \theta\right]$$
$$= J_0(K) \sin \theta + 2\sum_{n=1}^{\infty} J_n(K) \cos n\omega_a t \sin\left(\theta + \frac{n\pi}{2}\right) \quad (12)$$

and similarly for the terms containing the arguments $2K$ and $3K$. Omitting the constant terms in the frequency spectrum, the final result giving the nth-order distortion component becomes

$$f_p(n\omega) = \frac{4\beta\sqrt{q}}{(1+q)}\left\{\left[1 + \frac{1}{4}\left(\frac{1-q}{1+q}\right)^2\right] J_n(K) \sin\left(\theta + \frac{n\pi}{2}\right) - \frac{1}{2}\left(\frac{1-q}{1+q}\right) J_n(2K) \sin\left(2\theta + \frac{n\pi}{2}\right) + \frac{1}{4}\left(\frac{1-q}{1+q}\right)^2 J_n(3K) \sin\left(3\theta + \frac{n\pi}{2}\right)\right\} \cos n\omega_a t \quad (13)$$

β is, within a limited interval, approximately a linear function of the load v.s.w.r. Introducing the pulling figure p of the valve, which is one of the characteristics of a klystron and is defined as the change in frequency obtained when the phase of a load having a v.s.w.r. of 1·5 is changed through π radians, we have

$$\beta = \frac{S - 1}{1 \cdot 5 - 1} p = 2(S - 1)p$$

If the loss of the line is negligible, p is independent of the feeder length. The complete expression for the nth-order distortion component then becomes

$$f_p(n\omega) = \frac{8(S-1)p\sqrt{q}}{(1+q)}\left\{\left[1 + \frac{1}{4}\left(\frac{1-q}{1+q}\right)^2\right] J_n(K) \sin\left(\theta + \frac{n\pi}{2}\right) - \frac{1}{2}\left(\frac{1-q}{1+q}\right) J_n(2K) \sin\left(2\theta + \frac{n\pi}{2}\right) + \frac{1}{4}\left(\frac{1-q}{1+q}\right)^2 J_n(3K) \sin\left(3\theta + \frac{n\pi}{2}\right)\right\} \cos n\,\omega_a t \quad (14)$$

It should be noted that throughout the paper the feeder attenuation has been assumed negligible. If, therefore, there were a significant loss included in the line connecting the klystron to the mismatched load, it would be necessary to modify the parameters p and q in order to account for the reduced coupling between the modulator and the load. The modified (effective values of) p and q, however, are related very simply to the values obtained for a negligible feeder loss and will not be dealt with.

Eqn. (14) gives an exact value of the nth-order distortion component relative to the fundamental. As a result of the complexity of the expression, however, the evaluation becomes rather cumbersome and certain simplifications and the relevant approximations seem necessary and justifiable.

In practice the ratio ρ/ρ_s is fairly small (of the order of 0·2) and $\frac{1}{4}\left(\frac{1-q}{1+q}\right)^2$ in the above expression is negligible in comparison with unity. Similarly it can be shown that $\frac{1}{4}\left(\frac{1-q}{1+q}\right)^2 J_n(3K)$ is small in comparison with $J_n(K)$ for the components of interest, namely the fundamental, the second and the third. The term $\frac{1}{2}\left(\frac{1-q}{1+q}\right) J_n(2K)$, however, can be of the same order of magnitude as $J_n(K)$. In Table 1 the magnitudes of the various components of eqn. (14) are given for a typical value of K.

Table 1

COMPONENTS OF EQN. (14) FOR THE TYPICAL CASE $K = 0 \cdot 5$

n	$J_n(K)$	$\frac{1}{2}\left(\frac{1-q}{1+q}\right) J_n(2K)$	$\frac{1}{4}\left(\frac{1-q}{1+q}\right)^2 J_n(3K)$
1	0·24	0·044	0·005 6
2	0·03	0·011	0·002 3
3	0·002 5	0·002	0·000 6

$l = 10$ m, $f_d = 0 \cdot 8$ Mc/s, $f_c = 4\,000$ Mc/s, $\lambda g = 9 \cdot 5$ cm.
The ratio $\left(\frac{1-q}{1+q}\right)$ is taken as 0·2.

With these simplifications, the nth-order distortion component relative to the fundamental can be written

$$D_n = \frac{\frac{8(S-1)p\sqrt{q}}{1+q}\left[J_n(K) \sin\left(\theta + \frac{n\pi}{2}\right) - \frac{1}{2}\left(\frac{1-q}{1+q}\right) J_n(2K) \sin\left(2\theta + \frac{n\pi}{2}\right)\right]}{f_d + \frac{8(S-1)p\sqrt{q}}{1+q}\left[J_1(K) \cos \theta - \frac{1}{2}\left(\frac{1-q}{1+q}\right) J_1(2K) \cos 2\theta\right]} \quad (15)$$

In most modulator systems of practical significance, the fundamental distortion component is small compared with the undistorted component, as is clearly shown in the denominator of eqn. (15). Hence this equation reduces to

$$D_n = \frac{8(S-1)p\sqrt{q}}{(1+q)f_d}\left[J_n(K)\sin\left(\theta+\frac{n\pi}{2}\right) - \frac{1}{2}\left(\frac{1-q}{1+q}\right)J_n(2K)\sin\left(2\theta+\frac{n\pi}{2}\right)\right] \quad (15)$$

To obtain a suitable graphical representation of the results, the expression for D_n can be subjected to a certain degree of rationalization by writing

$$20\log D_n = 20\log\left\{\frac{8(S-1)\sqrt{q}}{(1+q)}\left[J_n(K)\sin\left(\theta+\frac{n\pi}{2}\right) - \frac{1}{2}\left(\frac{1-q}{1+q}\right)J_n(2K)\sin\left(2\theta+\frac{n\pi}{2}\right)\right]\right\} + 20\log\frac{p}{f_d} \quad (17)$$

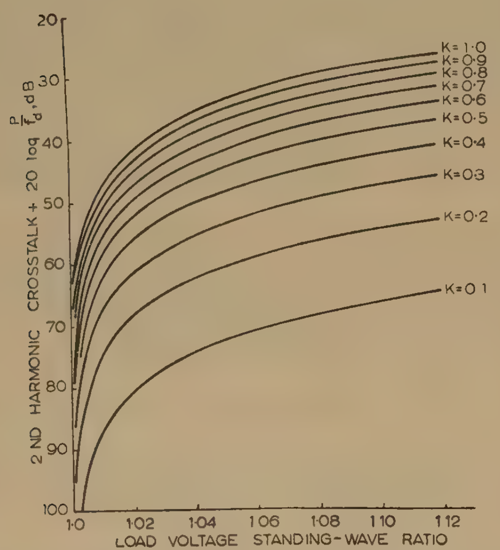


Fig. 4.—Second-harmonic crosstalk as a function of the load mismatch plotted with K as the parameter.

$\theta = 90°$, $S_s = 1\cdot3$.

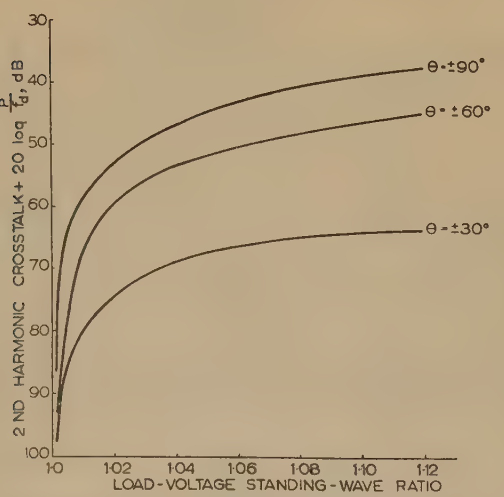


Fig. 5.—Second-harmonic crosstalk plotted to the base of v.s.w.r. with θ as the parameter.

$K = 0\cdot5$, $S_s = 1\cdot3$.

(3) A BRIEF DISCUSSION OF THE RESULTS

Figs. 4 and 5 give the second-harmonic crosstalk $+20\log$ (p plotted against the load v.s.w.r., the parameters being K θ respectively.

Fig. 4 shows that in the range of v.s.w.r. and delay time practical significance, the crosstalk tends to increase asympt cally towards a fixed value for increasing delay time. This fi value appears to be of the order of 25 dB less $20\log p/$ Practical values for the pulling figure p and the freque deviation f_d are 4 and 2 Mc/s, respectively, resulting in a limit value for the second-harmonic crosstalk of the order of 20 It should be noted that this value corresponds to the condit when θ, the effective angle of the load, is adjusted for maxim second-harmonic crosstalk and hence 20 dB refers to the w possible condition obtainable. Fig. 5 shows the effect of c trolling the load phase-angle θ. It is manifest that by adjust the angle θ to zero the second-harmonic crosstalk vanishes. superficial inspection of the typical Riecke diagram in Fig

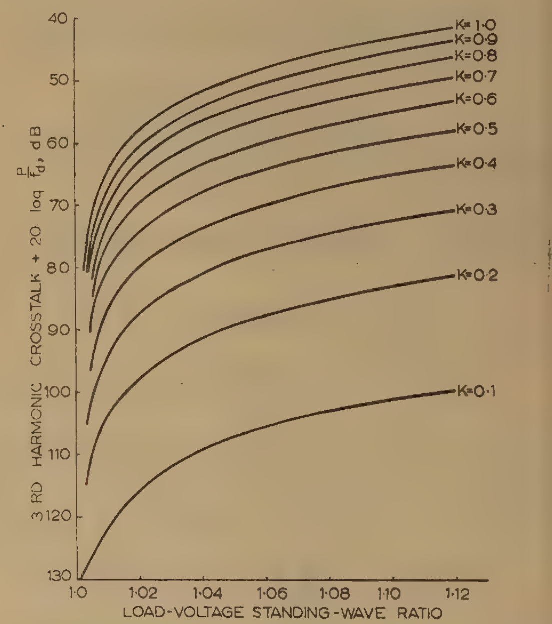


Fig. 6.—Third-harmonic crosstalk as a function of load mismat plotted with K as the parameter.

$\theta = 90°$, $S_s = 1\cdot3$.

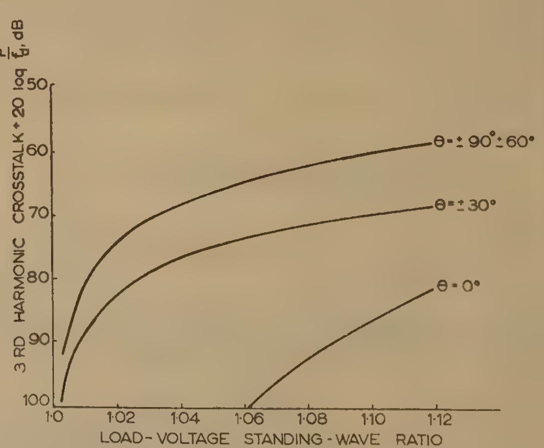


Fig. 7.—Third-harmonic crosstalk plotted to the base of v.s.w with θ as the parameter.

$K = 0\cdot5$, $S_s = 1\cdot3$.

reveals the symmetrical properties of the frequency characteristics of a klystron for varying load angles. On this basis alone one would anticipate the possibility of nullifying the second-harmonic distortion component and similarly all even harmonics.

A qualitative examination of the third-harmonic distortion component on the basis of the Riecke diagram is rather more complex as a result of the skewness of the lines constituting constant frequency. It does not, however, seem reasonable to expect the third-harmonic crosstalk to vanish for a certain value of the load phase angle. This is brought out in Figs. 6 and 7, where, as in Figs. 4 and 5, the third-order distortion components are plotted against the load v.s.w.r., θ and K being two independent parameters. It is shown that even the maximum third-harmonic crosstalk is far less severe than the second, but that it is not possible to neutralize the crosstalk completely by varying θ. It should be noted that by controlling the phase angle it is possible to minimize the total harmonic crosstalk, but it is in general impossible to cancel the cross-modulation completely.

(4) CONCLUSION

A simple theoretical investigation of the amount of crosstalk introduced by the pulling effect in a frequency-modulated klystron has been carried out. An expression giving the nth-order distortion component relative to the fundamental has been derived, and on the basis of this, introducing some simplifications accompanied by the inherent approximations, curves giving the second- and third-order distortion components relative to the fundamental have been computed. Little work has been done towards an experimental confirmation of the theoretical result obtained above, apart from a comparison of the results on a purely qualitative basis. A comprehensive experimental analysis of cross-modulation distortion arising from the pulling effect is somewhat cumbersome as a result of the number of distortion sources normally involved. Unless due precautions are taken to separate the different non-linear contributions in the modulator system, one is liable to measure the combined effect of 'echo distortion' and frequency-pulling distortion as well as that arising from the inherent non-linearity of the modulation characteristics of the klystron. By fairly simple means, however, one is able to confirm the theory qualitatively. The simple theoretical treatment of the problem of crosstalk arising from frequency pulling brings out the importance of isolating the antenna and feeder system from the modulated klystron, if the latter is working directly into the feeder, which frequently is the case in applications such as the back-to-back microlink system.

The computed curves show that even small reflections give rise to sizeable crosstalk if the feeder connecting the modulator to the load is not extremely short. Therefore, if the modulated klystron is not working directly into a power amplifier, it is advisable to connect a well-designed uniline directly to the modulator. From the point of view of crosstalk, the match of the load isolator need not then be very good.

It should again be noted that certain approximations were introduced in deriving the final expression for the distortion spectrum. It is therefore not advisable to apply the formula to cases where large standing-wave ratios are involved. These cases will in general be of little practical importance, since communication systems seldom operate with large mismatches on the feeders.

(5) ACKNOWLEDGMENTS

The author wishes to express his thanks to Mr. H. Sörby for his many helpful suggestions, and to Mr. K. Flöisand and Mr. J. Hannisdal for their assistance in computing the results. The paper is published by permission of the Director of the Norwegian Defence Research Establishment.

[The discussion on the above paper will be found on page 492.]

621.385.1.029.6

The Institution of Electrical Engineer
Paper No. 2788
Jan. 195

THEORY AND BEHAVIOUR OF HELIX STRUCTURES FOR A HIGH-POWER PULSED TRAVELLING-WAVE TUBE

By G. W. BUCKLEY, M.A., Associate Member, and J. GUNSON, B.A.

(*The paper was first received* 18*th July, and in revised form* 4*th October*, 1958. *It was published in January*, 1959, *and was read befo the* ELECTRONICS AND COMMUNICATIONS SECTION 23*rd February*, 1959.)

SUMMARY

The paper describes part of an investigation aimed at developing a multi-megawatt pulsed travelling-wave tube. To this end, a multi-start helix structure which may be water cooled has been considered both theoretically and experimentally.

Theoretical relationships have been determined for the axial series impedance of a sheath helix surrounded by a perfectly conducting cylinder, and for the phase velocity of the propagating modes. Results of experiments on four- and eight-start helix structures are found to give reasonable correspondence with the theoretical results for equivalent bounded sheath helices.

LIST OF SYMBOLS

ψ = Helix angle.
f = Frequency.
Δf = Perturbation frequency shift.
λ_0 = Free-space wavelength.
λ_g = Guide wavelength.
β_0 = Phase-change coefficient for transmission in free space.
β = Phase-change coefficient of helix structure.
c = Velocity of light.
v_1 = Phase velocity of travelling wave.
v_g = Group velocity of travelling wave.
γ = Radial propagation coefficient $= \beta\sqrt{\left[1 - \left(\frac{v_1}{c}\right)^2\right]}$
E_T = Peak value of axial electric field in travelling wave.
P = Total power flow in travelling wave.
v = Volume of perturbing object.
V_c = Volume of resonant cavity.
l = Length of resonant cavity.
a = Mean helix radius.
b = Bounding cylinder radius.

(1) INTRODUCTION

For many years the klystron was overshadowed by the magnetron as a source of high-level microwave power. However, the work of Chodorow *et al.*[1] has now established the klystron as a power generator, and for the greatest peak powers it has yet to be surpassed by a magnetron. Although, until fairly recently, the travelling-wave tube, with its great bandwidth, has occupied a role as a low-power source similar to the early klystron, it has now reached the megawatt level.[2]

With a view to developing a high-power travelling-wave tube, some overall performance parameters have been derived that might optimistically be expected from a valve obeying in every respect the simplest Pierce theory.[3] Analysis of an unshielded 3 cm-diameter helical sheath of 30° angle suggested a valve of about 30% efficiency, the corresponding gain being about 60 dB for 1 m length. As this elementary review showed promise, an introductory study was initiated into the general feasibility of a high-power pulsed travelling-wave tube, amplifying at a fre quency of 3 Gc/s and with a bandwidth of several hundre megacycles per second. Since the start of this investigatio added confirmation of the basic idea has come from the publishe results of Chodorow and Nalos,[2] and more recently, of Gitting *et al.*[4]

Practical experience with high-power pulsed klystrons sugges that the travelling-wave structure should be all metallic and co tained in a metal cylinder. To yield an r.f. power output o several megawatts, it seems reasonable to base an initial desig around the electron beam formed by the unity microperveanc gun used in 6 MW pulsed klystrons.[5] Such a gun gives a 200 k 90 amp electron beam which passes reasonably through a 2 cm diameter drift tube.

The foundations of a design having been laid by the choice o electron beam, it is necessary to consider the types of slow-wav structure that might give a reasonable bandwidth. These fa into two classes:

(*a*) The conventional helix.
(*b*) The periodic waveguide, typified by some variation of t iris-loaded circular waveguide used in linear electron accelerators

The helix-type travelling-wave tubes known at the present tim are relatively low-power devices, using an open structure forme by a single wire; they operate satisfactorily with electron bean formed by 2 kV 15 mA guns. However, it is quoted in t literature[2] that the wire helix is subject to power-handlin limitations which make it unsuitable for high-power application

Field[7] has made a survey of the properties of several loade waveguide structures. Analysis of his results suggests that the effective frequency bandwidth will be very narrow indeed co pared with the helix. Nevertheless, the approach suggested b Chodorow, Nalos[2] and Craig,[8] and particularly the idea of usi negative mutual coupling, leads to a 'linear accelerator' type loaded structure. The corresponding bandwidth is about 6 and the series impedance is high (about 150 ohms).

With such evidence available, it might be argued th Chodorow's approach is the one to adopt. However, as t main interest is in a travelling-wave tube with a bandwid greater than that yet obtained with the perforated-iris wavegui structure, it seems reasonable to re-examine the helix structu If it could be adapted for high-power operation its great ban width might be utilized, although its series impedance is kno to be low compared with the Chodorow structure. A way o of this difficulty has been suggested by Dodds and Peters,[9] w introduced the concept of the filter helix. Their technique to combine the helix with a periodic structure, thereby increasi the series impedance at the expense of the effective frequen bandwidth.

Most helix-type travelling-wave tubes have been designed following the techniques given by Pierce.[3] This design approa uses the concept of a helical sheath, to which a good approxim tion for the electron velocities normally used (i.e. up to one-ter the velocity of light) is a single-start wire helix. However, electron velocities of the order of two-thirds the velocity of lig

Mr. Buckley is, and Mr. Gunson was formerly, with the Metropolitan-Vickers Electrical Co. Ltd. Mr. Gunson is now at the University of Cambridge. At the London meeting the paper was read on their behalf by Dr. J. Brown.

a single-start helix of corresponding phase velocity is an open structure and is unsuitable for use in a high-voltage travelling-wave tube. Nevertheless, a high-power travelling-wave tube design may still be based on the sheath helix concept, and a multi-start helix is suggested as a practical form of structure for such a tube.

At first it may seem that the smallest feasible helix diameter is so large that the field interacting with the beam will be small. However, a closer study of the Pierce theory shows that the variation in the field across a 3 cm-diameter sheath of 30° angle is only about 25% at a frequency of 3 Gc/s.

Again, as high-power applications are envisaged, it is accepted that the helix will be water cooled. If this can be achieved in practice, the structure becomes ideal for high-power operation by virtue of the proximity of the coolant for transference of heat from both resistive losses and electron bombardment. The provision of a tubular helical conductor suggests a minimum bore for the water passage of $\frac{1}{8}$ in diameter which, combined with a wall thickness of $\frac{1}{32}$ in, leads to a minimum practical conductor cross-section of $\frac{3}{16}$ in diameter. Compared with the small-gauge wire conductor used hitherto as an approximation to the Pierce sheath, the large-section conductor is enormous. Consequently, the performance of a thick helical structure has to be assessed by experiment before its utility is known. Results so far obtained from such a study show sufficient promise to merit an alternative approach to that adopted by Chodorow.

(2) THE SHIELDED-HELIX SLOW-WAVE STRUCTURE

The unbounded helical sheath and tape helix structures have been reviewed extensively by Sensiper.[10] However, the bounded sheath helix (a helical sheath surrounded by a metal conducting cylinder) has only received scanty attention, and so the analysis has been extended and compared with experimental results derived by perturbation measurements.

Exact analysis of a single-start wire helix is difficult (see Sensiper[10]), and to set up a reasonable theory Pierce introduced the idea of the helically conducting sheet (or sheath helix). A helical sheath is represented in Fig. 1(a), where the cylinder is assumed to be infinitesimally thin in the radial direction. It is perfectly conducting in a helical direction making an angle ψ (the pitch angle) with a plane normal to the axis (the direction of propagation) and is non-conducting in a perpendicular direction to this ψ direction (the direction of conduction). Evaluation of the electrical characteristics follows from appropriate solutions of Maxwell's equations chosen for the inside and outside regions of the cylinder. Whilst the present study requires the solution for a bounded helical sheath, most of the published literature considers only an unbounded structure. Accordingly, the results for this latter structure are considered first.

(2.1) An Unbounded Helical Sheath

Pierce,[3] Chu and Jackson[11] have expressed the solutions in terms of three propagation parameters:

$$\left.\begin{aligned}\beta_0 &= \omega/c\\ \beta &= \omega/v_1\\ \gamma &= \beta\sqrt{\left[1-\left(\frac{v_1}{c}\right)^2\right]}\end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad (1)$$

where $\omega/2\pi$ is the frequency of the propagated wave.

Matching wave impedances in the usual way leads to a transcendental equation for the propagation factor, namely

$$\tan^2\psi = \left(\frac{\beta_0}{\gamma}\right)^2 \frac{I_1(\gamma a)K_1(\gamma a)}{I_0(\gamma a)K_0(\gamma a)} \quad . \quad . \quad . \quad (2)$$

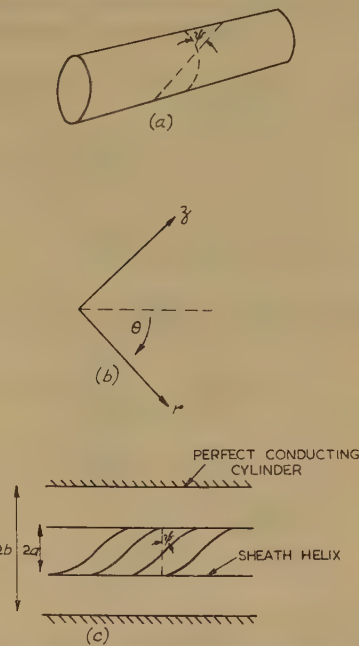


Fig. 1.—The helical sheath.

(a) Helical conducting sheet of radius a. The sheet is conducting along helical paths making an angle ψ with a plane normal to the axis.
(b) Right-hand set of polar co-ordinates.
(c) Schematic representation of bounded sheath helix (helix angle ψ).

where $I_n(x)$, $K_n(x)$ are modified Bessel functions of the first and second kind respectively and a is the radius of the helical sheath. M.K.S. units are used throughout.

From a consideration of the Poynting vector it follows that an associated axial series impedance parameter is given by the expression

$$\left(\frac{E_T^2}{\beta^2 P}\right)^{1/3} = \left(\frac{\beta}{\beta_0}\right)^{1/3}\left(\frac{\gamma}{\beta}\right)^{4/3} F(\gamma a) \quad . \quad . \quad . \quad (3)$$

where $F(\gamma a) = 7\cdot154\varepsilon^{-0\cdot664\gamma a}$

If $\frac{v_1}{c} \ll 1$ (which is not the case for a slow-wave structure associated with a 200 kV electron stream, for which $\frac{v_1}{c} \simeq 0\cdot67$), it follows from eqn. (2) that

$$\frac{v_1}{c} = \tan\psi \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

(2.2) The Bounded Helical Sheath

The structure of interest is a helical sheath of radius a surrounded by a perfect conductor of radius b. This configuration is shown schematically in Figs. 1(b) and (c) and the appropriate solution to Maxwell's equations may be derived from two Hertzian vectors. This technique was employed by Sensiper;[10] using the usual $\varepsilon^{-j\beta_n z}$ dependence, it leads to the solution

$$\Pi_n^i = \sum_n A_n^i I_n(\gamma r)\varepsilon^{-j\beta_n z}\varepsilon^{-jn\theta},\ a > r > 0 \quad . \quad . \quad . \quad . \quad (5)$$

$$\Pi_z^e = \sum_n [A_n^e K_n(\gamma r) + C_n^e I_n(\gamma r)]\varepsilon^{-j\beta_n z}\varepsilon^{-jn\theta},\ b > r > a \quad (6)$$

$$\Pi_n^{i*} = \sum_n B_n^i I_n(\gamma r)\varepsilon^{-j\beta_n z}\varepsilon^{-jn\theta},\ a > r > 0 \quad . \quad . \quad . \quad . \quad (7)$$

$$\Pi_z^{e*} = \sum_n [B_n^e K_n(\gamma r) + D_n^e I_n(\gamma r)]\varepsilon^{-j\beta_n z}\varepsilon^{-jn\theta},\ b > r > a \quad (8)$$

where β_n and γ have been defined in eqn. (1). A_n, B_n, C_n, D_n

are constants determined by the geometry of the system, and the indications i and e refer respectively to inside and outside the helical sheath.

In terms of two Hertzian vectors Π and Π^* the electromagnetic fields may be specified:

$$\left.\begin{aligned} E_r &= \frac{\partial^2 \Pi_z}{\partial z \partial r} - j\frac{\omega\mu}{r}\frac{\partial \Pi_z^*}{\partial \theta} \\ E_\theta &= \frac{1}{r}\frac{\partial^2 \Pi_z}{\partial z \partial \theta} + j\omega\mu\frac{\partial \Pi_z^*}{\partial r} \\ E_z &= \frac{\partial^2 \Pi_z}{\partial z^2} + \beta_0^2 \Pi_z \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad (9)$$

$$\left.\begin{aligned} H_r &= j\omega\epsilon\frac{1}{r}\frac{\partial \Pi_z}{\partial \theta} + \frac{\partial^2 \Pi_z^*}{\partial z \partial r} \\ H_\theta &= -j\omega\epsilon\frac{\partial \Pi_z}{\partial r} + \frac{1}{r}\frac{\partial^2 \Pi_z^*}{\partial z \partial \theta} \\ H_z &= \frac{\partial^2 \Pi_z^*}{\partial z^2} + \beta_0^2 \Pi_z^* \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad (10)$$

and $\beta_0^2 = \omega^2\mu\epsilon$ as in eqn. (1).

After a little algebra, it follows from eqns. (5)–(10) that the electromagnetic field components may be represented by infinite series, for which, the factor $\varepsilon^{-j\beta_n z}\varepsilon^{-jn\theta}$ being understood, the general terms may be written:

(i) Within the helical sheath ($0 < r < a$)

$$\left.\begin{aligned} E_r &= -j\beta\gamma A_n^i I_n'(\gamma r) - \frac{\omega\mu n}{r} B_n^i I_n(\gamma r) \\ E_\theta &= -\frac{\beta n}{r} A_n^i I_n(\gamma r) + j\omega\mu\gamma B_n^i I_n'(\gamma r) \\ E_z &= -\gamma^2 A_n^i I_n(\gamma r) \end{aligned}\right\} \quad . \quad . \quad (11)$$

$$\left.\begin{aligned} H_r &= \frac{\omega n\epsilon}{r} A_n^i I_n(\gamma r) - j\beta\gamma B_n^i I_n'(\gamma r) \\ H_\theta &= -j\omega\epsilon\gamma A_n^i I_n'(\gamma r) - \frac{\beta n}{r} B_n^i I_n(\gamma r) \\ H_z &= -\gamma^2 B_n^i I_n(\gamma r) \end{aligned}\right\} \quad . \quad . \quad (12)$$

(ii) Between the helical sheath and outer cylinder ($a < r < b$)

$$\left.\begin{aligned} E_r &= -j\beta\gamma[A_n^e K_n'(\gamma r) + C_n^e I_n'(\gamma r)] \\ &\quad - \frac{\omega\mu n}{r}[B_n^e K_n(\gamma r) + D_n^e I_n(\gamma r)] \\ E_\theta &= -\frac{\beta n}{r}[A_n^e K_n(\gamma r) + C_n^e I_n(\gamma r)] \\ &\quad + j\omega\mu\gamma[B_n^e K_n'(\gamma r) + D_n^e I_n'(\gamma r)] \\ E_z &= -\gamma^2[A_n^e K_n(\gamma r) + C_n^e I_n(\gamma r)] \end{aligned}\right\} \quad . \quad (13)$$

$$\left.\begin{aligned} H_r &= \frac{\omega\epsilon n}{r}[A_n^e K_n(\gamma r) + C_n^e I_n(\gamma r)] \\ &\quad - j\beta\gamma[B_n^e K_n'(\gamma r) + D_n^e I_n'(\gamma r)] \\ H_\theta &= -j\omega\epsilon\gamma[A_n^e K_n'(\gamma r) + C_n^e I_n'(\gamma r)] \\ &\quad - \frac{\beta n}{r}[B_n^e K_n(\gamma r) + D_n^e I_n(\gamma r)] \\ H_z &= -\gamma^2[B_n^e K_n(\gamma r) + D_n^e I_n(\gamma r)] \end{aligned}\right\} \quad . \quad (14)$$

To obtain the appropriate dispersion characteristics it is nece sary to match the fields indicated by eqns. (11)–(14) at the vario boundaries. The general method of field matching, which h been indicated by Slater,[12] is complex as it involves infini series. However, in the case of the sheath helix the bounda conditions are the same for all values of θ and z and so t conditions may be satisfied for each value of n separately.

The boundary conditions are examined at $r = a$ and $r =$ [The case of $r = 0$ has been included in the form of the solutio taken for eqns. (5) and (7).]

When $r = b$, the electric field must be normal to the conducti surface and so

$$E_z = E_\theta = 0$$

When $r = a$, the nature of the helical conductivity is shown the relations

$$E_z^i = E_z^e \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1$$

$$E_\theta^i = E^e \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1$$

$$E_z^i = -E_\theta^i \cot\psi \quad . \quad . \quad . \quad . \quad (1$$

$$E_z^e = -E_\theta^e \cot\psi \quad . \quad . \quad . \quad . \quad (1$$

$$(H_z^i - H_z^e) = (H_\theta^e - H_\theta^i)\cot\psi \quad . \quad . \quad . \quad (1$$

and so by suitable manipulation

$$C_n^e = -A_n^e \frac{K_n(\gamma b)}{I_n(\gamma b)} \quad . \quad . \quad . \quad . \quad . \quad . \quad (2$$

$$D_n^e = -B_n^e \frac{K_n'(\gamma b)}{I_n'(\gamma b)} \quad . \quad . \quad . \quad . \quad . \quad . \quad (2$$

$$A_n^i = A_n^e \frac{K_n(\gamma a)}{I_n(\gamma a)}\left[1 - \frac{I_n(\gamma a)K_n(\gamma b)}{K_n(\gamma a)I_n(\gamma b)}\right] \quad . \quad . \quad . \quad . \quad . \quad . \quad (2$$

$$B_n^i = A_n^e \frac{K_n(\gamma a)}{j\omega\mu\gamma I_n'(\gamma a)}\left(\frac{\beta n}{a} + \gamma^2 \tan\psi\right)\left[1 - \frac{I_n(\gamma a)K_n(\gamma b)}{K_n(\gamma a)I_n(\gamma b)}\right] \quad (2$$

$$B_n^e = A_n^e \frac{K_n(\gamma a)}{j\omega\mu\gamma K_n'(\gamma a)} \frac{1 - \dfrac{I_n(\gamma a)K_n(\gamma b)}{I_n(\gamma b)K_n(\gamma a)}}{1 - \dfrac{I_n'(\gamma a)K_n'(\gamma b)}{I_n'(\gamma b)K_n'(\gamma a)}} \quad . \quad . \quad . \quad . \quad (2$$

Thus, after some algebra,

$$\left[\frac{1 - \dfrac{I_n'(\gamma a)K_n'(\gamma b)}{I_n'(\gamma b)K_n'(\gamma a)}}{1 - \dfrac{I_n(\gamma a)K_n(\gamma b)}{I_n(\gamma b)K_n(\gamma a)}}\right] \frac{I_n'(\gamma a)K_n'(\gamma a)}{I_n(\gamma a)K_n(\gamma a)} = -\frac{\left(\dfrac{\beta n}{a} + \gamma^2 \tan\psi\right)^2}{\gamma^2\beta_0^2} \quad (2$$

For the case of the unbound helix, Sensiper[10] derived the cor sponding expression

$$\frac{I_n'(\gamma a)K_n'(\gamma a)}{I_n(\gamma a)K_n(\gamma a)} = -\frac{\left(\dfrac{\beta n}{a} + \gamma^2 \tan\psi\right)^2}{\gamma^2\beta_0^2} \quad . \quad . \quad (2$$

The only non-zero axial electric-field component correspo to $n = 0$, and with this substitution, eqn. (25) becomes

$$\tan^2\psi = \left(\frac{\beta_0}{\gamma}\right)^2 \frac{I_1(\gamma a)K_1(\gamma a)}{I_0(\gamma a)K_0(\gamma a)}\left[\frac{1 - \dfrac{I_1(\gamma a)K_1(\gamma b)}{K_1(\gamma a)I_1(\gamma b)}}{1 - \dfrac{I_0(\gamma a)K_0(\gamma b)}{K_0(\gamma a)I_0(\gamma b)}}\right] \quad ($$

Eqn. (27) reduces to eqn. (2), the corresponding equation the helix in free space, as b becomes large.

The properties of the function

$$G(x) = \frac{1 - \dfrac{I_1(x)K_1(nx)}{K_1(x)I_1(nx)}}{1 - \dfrac{I_0(x)K_0(nx)}{K_0(x)I_0(nx)}} \quad . \quad . \quad . \quad . \quad (28)$$

are illustrated in Fig. 2, and it may be seen that for $x \geqslant 1$ and for $n > 1\cdot5$, $G(x) \to 1$. This property is convenient, for under these conditions the simple Pierce result, eqn. (2), may be used to determine the dispersion characteristic.

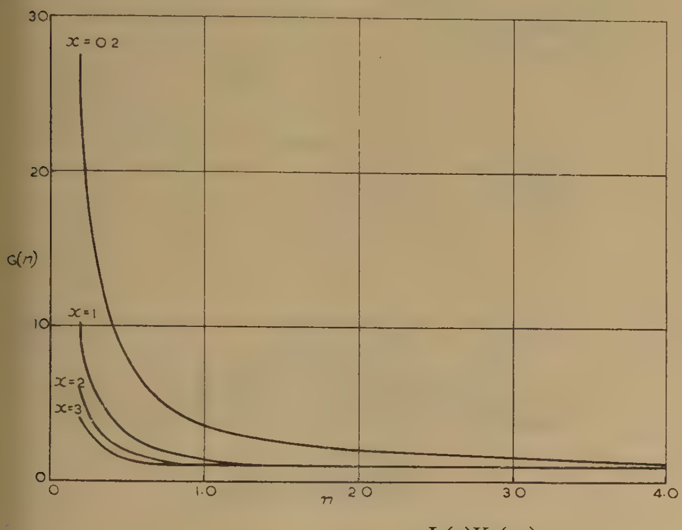


Fig. 2.—Plot of G(n) $= \dfrac{1 - \dfrac{I_1(x)K_1(nx)}{K_1(x)I_1(nx)}}{1 - \dfrac{I_0(x)K_0(nx)}{K_0(x)I_0(nx)}}$

The calculation of the axial series impedance defined as E^2/β^2P is quite involved. E_z is known, whilst β follows from the dispersion characteristic. Use is made of the Poynting vector in the evaluation of the total axial power flow, P, and it is helpful to separate P into two main parts so that $P = P_1 + P_2$.

Before evaluating P, it is convenient to rewrite eqns. (11)–(14) for the case of $n = 0$ in the form

$$\left.\begin{aligned}
H_z &= AI_0(\gamma r) + BK_0(\gamma r)\\
H_r &= \frac{j\beta}{\gamma}[AI_1(\gamma r) - BK_1(\gamma r)]\\
E_\theta &= -\frac{j\omega\mu}{\gamma}[AI_1(\gamma r) - BK_1(\gamma r)]\\
E_z &= CI_0(\gamma r) + DK_0(\gamma r)\\
E_r &= \frac{j\beta}{\gamma}[CI_1(\gamma r) - DK_1(\gamma r)]\\
H_\theta &= \frac{j\omega\epsilon}{\gamma}[CI_1(\gamma r) - DK_1(\gamma r)]
\end{aligned}\right\} \quad . \quad . \quad (29)$$

the additional notation to be employed being:

(i) Within the helical conducting cylinder, i.e. $0 \leqslant r \leqslant a$, the suffix 1 is used.

(ii) Between the helical conducting cylinder and the outer conducting sheath, i.e. $a \leqslant r \leqslant b$, the suffix 2 is used.

Accordingly,

$$B_1 = D_1 = 0$$

$$B_2 = A_2\frac{I_1(\gamma b)}{K_1(\gamma b)}$$

$$D_2 = -C_2\frac{I_0(\gamma b)}{K_0(\gamma b)}$$

$$C_1 = \frac{j\omega\mu}{\gamma}A_2\left[\frac{I_1(\gamma a)K_1(\gamma b) - I_1(\gamma b)K_1(\gamma a)}{I_0(\gamma a)K_0(\gamma b)}\right]\cot\psi$$

$$C_2 = \frac{j\omega\mu}{\gamma}A_2\frac{K_0(\gamma b)}{K_1(\gamma b)}\left[\frac{I_1(\gamma a)K_1(\gamma b) - I_1(\gamma b)K_1(\gamma a)}{I_0(\gamma a)K_0(\gamma b) - I_0(\gamma b)K_0(\gamma a)}\right]\cot\psi$$

$$A_1 = A_2\left[\frac{I_1(\gamma a)K_1(\gamma b) - I_1(\gamma b)K_1(\gamma a)}{I_1(\gamma a)K_1(\gamma b)}\right]$$

It follows that on the axis $E_z = C_1$.

In general, if H^* is the complex conjugate of H,

$$P_z = 2\pi\int_{\text{cross-section}} \tfrac{1}{2}\mathscr{R}(E_rH_\theta^* - E_\theta H_r^*)rdr$$

so that, inside the helical sheath $(0 < r < a)$,

$$\frac{P_1}{|C_1|^2} = \frac{\pi}{|C_1|^2}\int_0^a \frac{\beta\omega}{\gamma^2}(\epsilon|C_1|^2 + \mu|A_1|^2)I_1^2(\gamma a)rdr$$

$$= \frac{\pi a^2\beta\omega}{2\gamma^2}\left[\epsilon + \frac{\gamma^2}{\mu\omega^2}\frac{I_0^2(\gamma a)}{I_1^2(\gamma a)}\tan^2\psi\right] \times \left[I_1^2(\gamma a) + \frac{2I_1(\gamma a)I_0(\gamma a)}{\gamma a} - I_0^2(\gamma a)\right]$$

and outside the helix $(a \leqslant r \leqslant b)$,

$$P_2 = \pi\left[\frac{\omega\mu\beta}{\gamma^2}|A_2|^2 + \frac{\omega\epsilon\beta}{\gamma^2}|C_2|^2\right]\int_a^b I_1^2(\gamma r)rdr$$
$$-2\pi\left[\frac{\omega\mu\beta}{\gamma^2}\frac{I_1(\gamma b)}{K_1(\gamma b)}|A_2|^2 - \frac{\omega\epsilon\beta}{\gamma^2}\frac{I_0(\gamma b)}{K_0(\gamma b)}|C_2|^2\right]\int_a^b I_1(\gamma r)K_1(\gamma r)rdr$$
$$+\pi\left[\frac{\omega\mu\beta}{\gamma^2}\frac{I_1^2(\gamma b)}{K_1^2(\gamma b)}|A_2|^2 + \frac{\omega\epsilon\beta}{\gamma^2}\frac{I_0^2(\gamma b)}{K_0^2(\gamma b)}|C_2|^2\right]\int_a^b K_1^2(\gamma r)rdr$$

In view of the length and complexity of the function representing P_2, the presentation is simplified by introducing the quantities P_α, P_β and P_γ, where $P_2 = P_\alpha + P_\beta + P_\gamma$ and

$$\frac{P_\alpha}{|C_1|^2} = \frac{\pi\omega\beta}{2\gamma^2}I_0^2(\gamma a)\left\{\frac{\epsilon K_0^2(\gamma b)}{[I_0(\gamma a)K_0(\gamma b) - I_0(\gamma b)K_0(\gamma a)]^2} + \frac{\gamma^2K_1^2(\gamma b)\tan^2\psi}{\omega^2\mu[I_1(\gamma a)K_1(\gamma b) - I_1(\gamma b)K_1(\gamma a)]^2}\right\}$$
$$\times\left\{[b^2I_1^2(\gamma b) - a^2I_1^2(\gamma a)] - [b^2I_0^2(\gamma b) - a^2I_0^2(\gamma a)] + \frac{2}{\gamma}[bI_1(\gamma b)I_0(\gamma b) - aI_1(\gamma a)I_0(\gamma b)]\right\}$$

$$\frac{P_\beta}{|C_1|^2} = \frac{\pi\omega\beta}{\gamma^2}I_0^2(\gamma a)\left\{\frac{\epsilon K_0(\gamma b)I_0(\gamma b)}{[I_0(\gamma a)K_0(\gamma b) - I_0(\gamma b)K_0(\gamma a)]^2} - \frac{\gamma^2}{\omega^2\mu}\frac{I_1(\gamma b)K_1(\gamma b)\tan^2\psi}{[I_1(\gamma a)K_1(\gamma b) - I_1(\gamma b)K_1(\gamma a)]^2}\right\}$$

$$\times \left\{ a^2 \left[\frac{I_1(\gamma a)K_0(\gamma a)}{\gamma a} - \frac{I_0(\gamma a)K_1(\gamma a)}{\gamma a} - I_0(\gamma a)K_0(\gamma a) - I_1(\gamma a)K_1(\gamma a) \right] \right.$$

$$\left. - b^2 \left[\frac{I_1(\gamma b)K_0(\gamma b)}{\gamma b} - \frac{I_0(\gamma b)K(\gamma b)}{\gamma b} - I_0(\gamma b)K_0(\gamma b) - I_1(\gamma b)K_1(\gamma b) \right] \right\}$$

$$\frac{P_\gamma}{|C_1|^2} = \frac{\pi\omega\beta}{2\gamma^2} I_0^2(\gamma a) \left\{ \frac{\epsilon I_0^2(\gamma b)}{[I_0(\gamma a)K_0(\gamma b) - I_0(\gamma b)K_0(\gamma a)]^2} + \frac{\gamma^2}{\omega^2\mu} \frac{I_1^2(\gamma b)\tan^2\psi}{[I_1(\gamma a)K_1(\gamma b) - I_1(\gamma b)K_1(\gamma a)]^2} \right\}$$

$$\times \left\{ a^2 \left[K_0^2(\gamma a) + \frac{2K_1(\gamma a)K_0(\gamma a)}{\gamma a} - K_1^2(\gamma a) \right] - b^2 \left[K_0^2(\gamma b) + \frac{2K_1(\gamma b)K_0(\gamma b)}{\gamma b} - K_1^2(\gamma b) \right] \right\}$$

Fig. 3A shows the variation of phase velocity for a 3 cm-diameter unbounded sheath helix as determined by computation

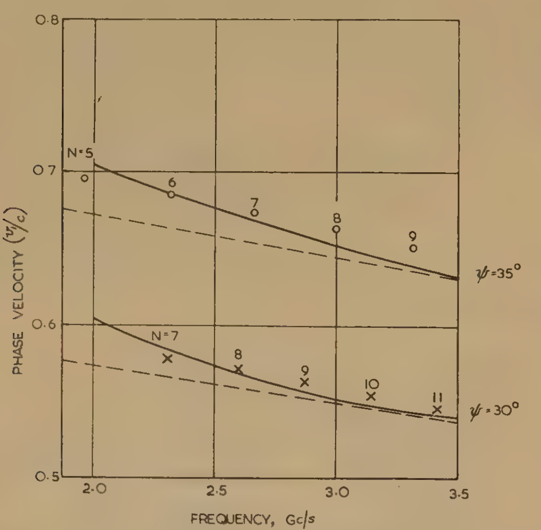


Fig. 3A.—Theoretical and experimental dispersion curves for sheath helix (fundamental mode).

—— Theoretical curve, bounded sheath helix.
---- Theoretical curve, unbounded sheath helix.
× Experimental points, 4-start helix.
○ Experimental points, 8-start helix.

from eqn. (2); the broadband nature of the structure is at once apparent. For synchronism of the wave with a 200 kV electron beam the helix angle is about 30°, the corresponding axial series impedance being 33 ohms.

The presence of a surrounding conducting cylinder modifies this dispersion characteristic in the manner indicated by eqn. (27). For a helical sheath diameter of 3 cm, a bounding cylinder of 6 cm diameter was chosen. Reference to Fig. 2 shows that the corresponding dispersion characteristic should approximate closely to the corresponding unbounded characteristic seen in Fig. 3A.

Calculation of the series impedance of this shielded helix structure requires an assessment of the axial power flow. Fig. 4 shows this characteristic for a 30° helix angle, with the relative powers flowing inside and outside the helical sheath. When combined with the theoretical dispersion characteristics of

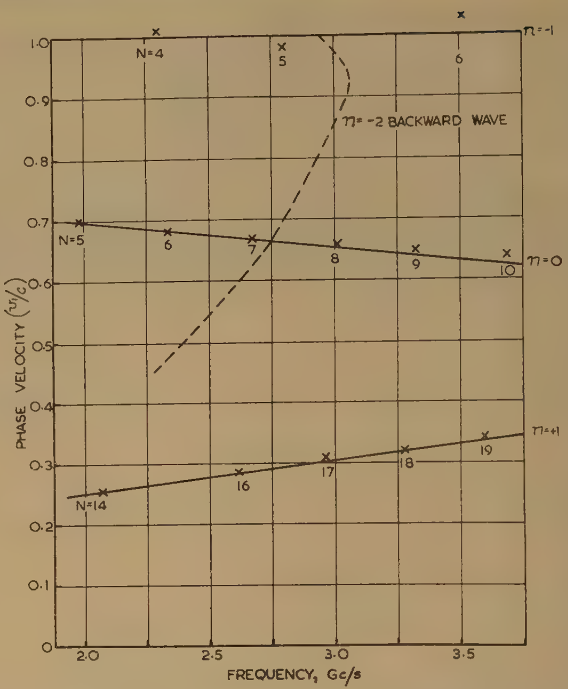


Fig. 3B.—Comparison of experimental resonances found for unstrapped helix, with bounded sheath helix theory.

---- Theoretical curves.
× Experimental points.
N Resonant mode numbers.

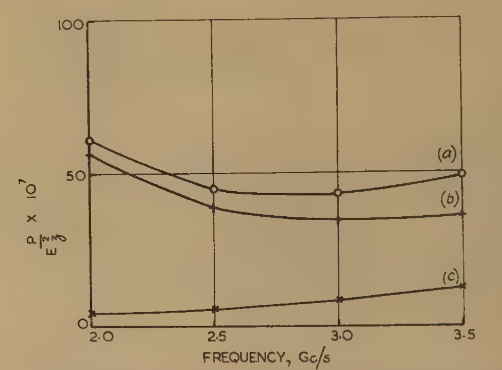


Fig. 4.—Normalized axial power flow in bounded sheath helix.

Sheath helix diameter, 3 cm. Outer conductor diameter, 6 cm. Helix angle, 30°
(a) Total power flow characteristic.
(b) Power flow outside helical sheath.
(c) Power flow inside helical sheath.

Fig. 3A, the variation of axial series impedance with frequency represented in Fig. 5.

It is instructive to compare this impedance characteristic w the Pierce expression of eqn. (3). The parameter $F(\gamma a)$ for unbounded helix of 3 cm diameter may be compared with equivalent function computed for the same helix surrounded coaxial conducting cylinders of different diameters. Fig. shows that a 2 : 1 diameter ratio does not introduce too serio a reduction in the axial series impedance, thereby confirming choice of basic diameters.

(3) EXPERIMENTAL METHODS

The two important characteristics are the phase velocity a the axial series impedance for the fundamental mode. For structures considered, these were determined by resonant-f

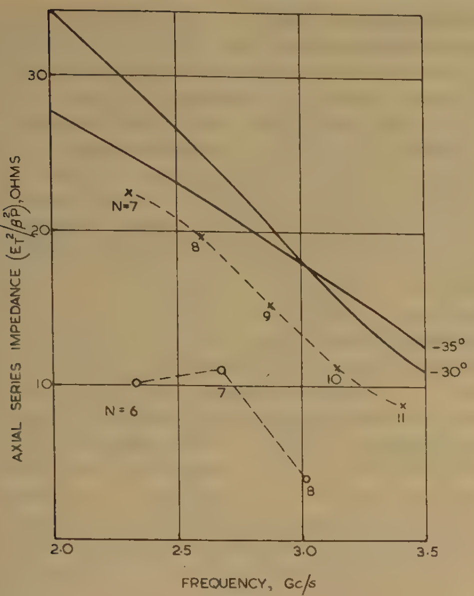


Fig. 5.—Theoretical and experimental series impedance for bounded sheath helix.

—— Theoretical curves, bounded sheath helix.
× Experimental points, 4-start 30° helix.
○ Experimental points, 8-start 35° helix.

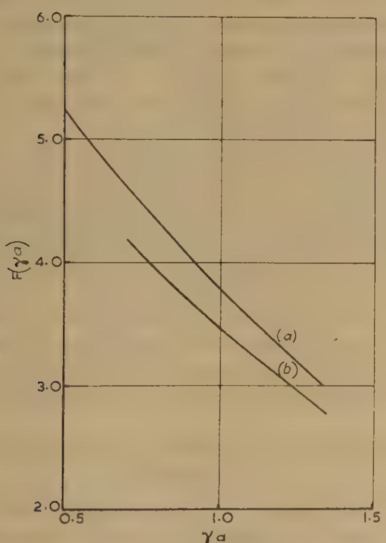


Fig. 6.—Plot of $F(\gamma a) = \dfrac{\left(\dfrac{E^2}{\beta^2 P}\right)^{1/3}}{\left(\dfrac{\beta}{\beta_0}\right)^{1/3}\left(\dfrac{\gamma}{\beta}\right)^{4/3}}$ for 35° helix.

(a) Unbounded sheath helix.
(b) Bounded sheath helix. (Outer-diameter/helix-sheath-diameter = 2.)

quency and perturbation measurements on cavities formed by enclosing suitable lengths of the structure between short-circuiting planes (Fig. 7).

(3.1) Phase Velocity Determination

With ideal short-circuits, the cavity would be an integral number of half-wavelengths long at resonance. The helix has no planes of symmetry, so in a standing-wave pattern there are no planes where a metallic plate can be inserted without disturbing the field distribution. Higher evanescent modes are excited near the ends of the cavity and modify the measured field-strength patterns. However, it has been shown[13] that, if there are no other circularly symmetric propagating modes of the helix for the range of parameters considered, a metallic plane short-circuit is equivalent to an ideal short-circuit somewhat displaced from the physical plane.

Fig. 7.—Strapped eight-start helix.
Half outer shield removed.

The plots of axial electric field strength using a perturbing bead are described below and serve to determine approximately the number of half-wavelengths, N, in the length of the cavity. Nevertheless, the guide wavelength may be determined to sufficient accuracy by assuming the cavity to be exactly an integral number of half-wavelengths long, so that

$$\frac{v_1}{c} = \frac{2L}{N}\frac{f}{c} \quad \ldots \ldots \quad (30)$$

(3.2) Axial Series Impedance Determination

Since both longitudinal magnetic and electric fields are present along the axis of the helix, a dielectric bead was chosen in order to perturb only the electric component. The relevant theory of cavity perturbation by a small object has been fully described by several authors.[14,15]

Applied to a small dielectric sphere of volume v and dielectric constant ϵ, it gives

$$\frac{E_T^2}{\beta^2 P} = -\frac{2}{3v}\left(\frac{\epsilon+2}{\epsilon-1}\right)\frac{L}{\epsilon_0 v_g}\frac{\Delta f}{f} \quad \ldots \quad (31)$$

The presence of higher modes near the ends of the cavity invalidates any measurements of series impedance in these regions. One effect of these modes can be seen in Fig. 8, where there is up to 10% variation in the peak electric field strengths. The group velocity v_g, is obtained by differentiating a dispersion curve for the structure.

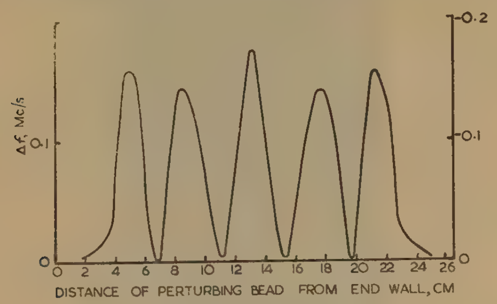


Fig. 8.—Typical perturbation curve.
Frequency shift, Δf in Mc/s.

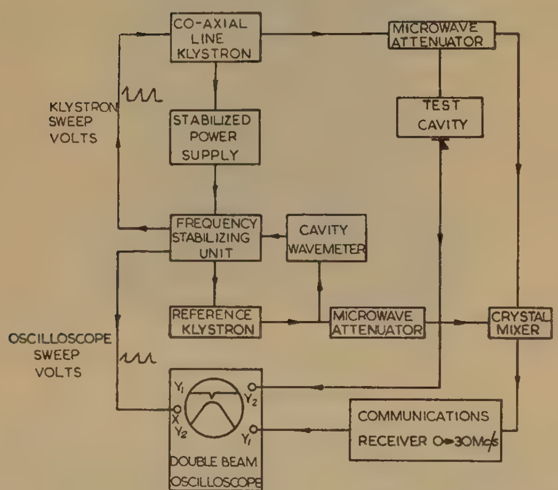


Fig. 9.—Arrangement of measuring equipment.

(4) EXPERIMENTAL DETAILS

The schematic of the apparatus used is shown in Fig. 9. A coaxial-line klystron oscillator generated at frequencies from 2 to 4 Gc/s, and the main resonant frequencies of the cavity were measured by comparison with a number of cavity wavemeters. A swept-frequency technique was used, and the resonance curves were displayed directly on an oscilloscope screen. The test cavity, Fig. 10, was monitored by a small loop in an end wall,

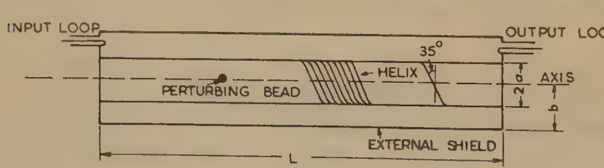


Fig. 10.—Diagram of test cavity.

oriented so as to couple with the circumferential magnetic field of the fundamental mode.

A small glass sphere suspended by a fine nylon thread along the axis of the test cavity was used to perturb the resonant frequency. The field variation along the axis was determined by plotting the frequency deviation Δf against the distance of th bead from one end of the cavity. Ideally, the result should be (cosine)^2 curve, but as a result of higher modes, it is distorte especially near the ends of the cavity. By assuming the vari tions in Δf to be random, an estimate of the actual frequenc shift was obtained by averaging the peak values of Δf at positio not too close to the ends of the cavity.

In some cases the measured values of Δf were of the order 200 kc/s, and therefore not much greater than the frequen discrimination of the wavemeters, which was ± 30 kc/s over mo of the frequency range, due to reading error. To obtain mo accurate values, the cavity wavemeter was replaced by a heter dyne frequency-measuring system. This used a second klystr oscillator operating at a fixed frequency determined by a cavi wavemeter. A difference frequency of a few megacycles p second was produced when power from both klystrons was fe into a crystal mixer, the mixer output being fed into a standa communication receiver. An output signal was taken from th receiver and displayed as a narrow pip on the oscilloscope scree thereby acting as the 'wavemeter' trace. The change in resona frequency caused by perturbation of the test cavity was measure directly from the frequency calibration of the communicatio receiver. With an effective Q-factor of about half a million, th discrimination was better than ± 5 kc/s. This compared wi the accuracy of setting the resonance peaks into coincidenc which was at best ± 10 kc/s.

Measurements were complicated by the presence of numero resonances when the cavity oscillated in other, non-circular symmetric, modes. Such resonances were distinguished fro true fundamental-mode resonances by the absence of true zer of electric field along the axis. When the guide wavelength w not too short, perturbation plots of some of these resonanc could be obtained, from which the phase velocities could b determined by counting the peaks.

To investigate the effect of strapping on the electrical chara teristics, thin copper bands, $\frac{1}{8}$ in wide, were clamped round th helix (Fig. 7). These were spaced at equal distances along th helix, the spacing being such that the short-circuit planes alwa coincided with a symmetry plane of the set of straps. For ea in assembly this was always chosen to be half-way between tw straps. The measurements on strapped helices were easier sin resonances due to higher modes tended to be less prominen this effect presumably being associated with the large perturbi effect that the straps had on the transverse modes.

(5) EXPERIMENTAL RESULTS. COMPARISON WITH THEORY

The most complete set of measurements were made on th eight-start 35° helix, this being the original structure chosen aft calculations with sheath helix theory. Phase velocity and axi series impedance curves are shown in Figs. 3A, 3B and 5. T experimental values of phase velocity agree reasonably well wi theory, though the deviations shown are significantly great than experimental error. Cavity lengths and frequencies a measurable to better than 1 part in 1 000, and so the accuracy measuring phase velocity is of the same order. The deviatio are attributed to two main causes:

(*a*) The assumption that the cavity length is exactly an integ number of half-wavelengths long.

(*b*) Mechanical imperfections in the helix due to random devi tions from the design values of conductor spacing and radii structures.

A more sensitive test of the difference between multi-start a sheath helices is given by series impedance measuremen Fig. 5 shows that agreement between experiment and theory closer for the four-start helix, the experimental curve being alm

the same as the theoretical curve except for a constant reduction factor of about 0·75. The experimental errors are not shown on the diagram, but are of the order ±5% for the four-start helix, and ±20% for the eight-start helix. For the latter, the measured values fall well below the theoretical curve, showing that the ratio of conductor thickness to conductor separation is too high for sufficient coupling between the inner and outer regions of the structure. In addition, slight variations in the pitch of separate conductors cause large percentage variations in the gap width, which is nominally 0·07 in. On the other hand, the gap width for the four-start helix is 0·4 in, and so not only are these variations reduced, but the coupling is increased. Obviously there must be an optimum number of conductors for a given helix angle and conductor diameter, to give greatest series impedance in the fundamental mode.

A knowledge of higher propagating modes is of value in determining the possibility of unwanted coupling to an electron beam. Such interaction could lead to instability in travelling-wave-tube operation. Accordingly, theoretical calculations of the phase velocities of these higher modes has been carried out using eqn. (25). The results are plotted in Fig. 3B for the 35° helix, with corresponding experimental results. Some of the experimental phase velocity results were derived from measurements of appropriate standing-wave patterns. The other experimental values have been deduced from cavity resonances when the corresponding standing-wave pattern defied resolution. Johnson, Everhart and Seigman[16] show that the space harmonic content of the fundamental mode on an N-start helix contains only components corresponding to the . . . $-2N$, $-N$, O, N, $2N$, . . . modes on a sheath helix. For the four-start and eight-start structures, the phase velocities are well removed from the region of interest. The only mode which might cause instability is the $n = -2$ backward-wave mode shown in Fig. 3B. Interaction with this mode would be strong only for parts of the electron beam well off the axis and near the helix structure, because of the $I_2(\gamma r)$ variation in the longitudinal electric field strength. Satisfactory determination of the effects of these modes can only be made by actual electron beam measurements.

Determination of the effects of strapping were carried out only

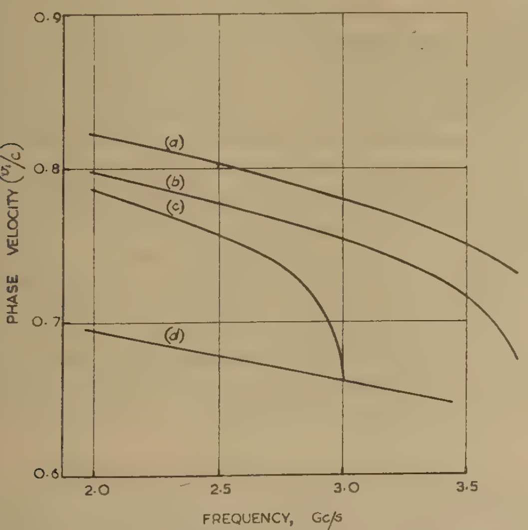


Fig. 11.—Strapped helix dispersion characteristics.

35° 8-start helix.

Strap spacings
(a) 2·20 cm.
(b) 2·64 cm.
(c) 3·30 cm.
(d) No straps.

on the eight-start helix. The phase velocity curves in Fig. 11 show a general increase in value, except near the π-mode region, where the group velocity falls to zero and the phase velocity approaches the value for the unstrapped helix curve. The measurements of series impedance are not very accurate, but the general trend shows that it is of the same order as for the unstrapped structure, increasing only when the group velocity falls near the π-mode cut-off. In order to obtain an appreciable increase in series impedance, heavier loading is necessary, possibly internal to the helix. This would increase both the relative axial electric field values and reduce the group velocity.

(6) FUTURE DEVELOPMENTS

The success of the bounded sheath theory in predicting the dispersion relationship (Fig. 3B) for a practical structure is gratifying.

Helices so far measured experimentally have been formed from separate conductors. However, the small effect of the straps noted in the measurements allows such members to be present in an advanced structure. Accordingly, a unit construction system may be developed using standard 'bricks' of the type shown in Fig. 12.

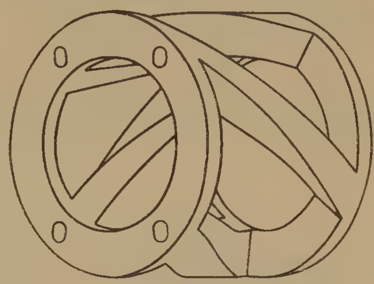


Fig. 12.—Unit brick.

An assembly technique to simulate the unidirectional helix structure is obvious. However, if adjacent elements are assembled to give alternately right-hand and left-hand helix sections, a simple theory suggests that the resultant pseudo-cross-wound structure would have a greater series impedance than the corresponding continuous helix.

(7) ACKNOWLEDGMENT

The authors thank Sir Willis Jackson, F.R.S., Director of Research and Education, Metropolitan-Vickers Electrical Co. Ltd., for permission to publish the paper.

(8) REFERENCES

(1) Chodorow, M., Ginzton, E. L., Neilson, I. R., and Sonkin, S.: 'Design and Performance of a High-Power Pulsed Klystron', *Proceedings of the Institute of Radio Engineers*, 1953, **41**, p. 1584.

(2) Chodorow, M., and Nalos, E. J.: 'The Design of High-Power Traveling-Wave Tubes', *ibid.*, 1956, **44**, p. 649.

(3) Pierce, J. R.: 'Traveling-Wave Tubes' (Van Nostrand, 1950).

(4) Gittings, J. F., Rock, N. H., and Sullivan, A. B. J.: 'An Experimental High-Power Pulsed Travelling Wave Tube', *Services Electronics Research Laboratory Technical Journal*, 1957, **7**, p. 42.

(5) Thomson, J.: 'Developments in Thermionic-Devices since 1939', *Journal I.E.E.*, 1955, **1** (New Series), p. 160.

(6) Mullett, L. B., and Loach, B. G.: 'Experimental Work on Corrugated Wave-Guides and Associated Components

for Linear Accelerators', *Proceedings of the Physical Society*, 1948, **61**, p. 271.

(7) Field, L. M.: 'Some Slow-Wave Structures for Traveling Wave Tubes', *Proceedings of the Institute of Radio Engineers*, 1949, **37**, p. 34.

(8) Chodorow, M., and Craig, R. A.: 'Some New Circuits for High-Power Traveling Wave Tubes', *ibid.*, 1957, **45**, p. 1106.

(9) Dodds, W. J., and Peters, R. W.: 'Filter-Helix Traveling Wave Tube', *RCA Review*, 1953, p, 502.

(10) Sensiper, S.: 'Electromagnetic Wave Propagation on Helical Conductors', D.Sc. Thesis, Massachusetts Institute of Technology, 1951.

(11) Chu, L. J., and Jackson, J. D.: 'Field Theory of Traveling Wave Tubes', *Proceedings of the Institute of Radio Engineers*, 1948, **36**, p. 853.

(12) Slater, J. C.: 'The Design of Linear Accelerato Massachusetts Institute of Technology Report No. 1947.

(13) Watson, W. H., and Whinnery, J. R.: 'Study of a Pla Short on a Shielded Helix', *Transactions of the Instit of Radio Engineers, Professional Group on Electr Devices*, 1955, **2**, p. 34.

(14) Casimir, H. B. G.: 'On the Theory of Electromagne Waves in Resonant Cavities', *Philips Research Repor* 1951, **6**, p. 162.

(15) Aikin, A. W.: 'Measurements in Travelling Wave Stru tures', *Wireless Engineer*, 1955, **32**, p. 230.

(16) Johnson, H. R., Everhart, T. E., and Seigman, A. 'Wave Propagation on Multifilar Helices', *Transactio of the Institute of Radio Engineers, Professional Group Electron Devices*, 1956, **3**, p. 18.

[The discussion on the above paper will be found on page 492.]

621.385.1.029.6

The Institution of Electrical Engineers
Paper No. 2825 E
Jan. 1959

A MULTI-CAVITY KLYSTRON WITH DOUBLE-TUNED OUTPUT CIRCUIT

By H. J. CURNOW, B.Sc., and L. E. S. MATHIAS, M.Sc.

(*The paper was first received* 17*th September, and in revised form* 13*th November*, 1958. *It was published in January*, 1959, *and was read before the* ELECTRONICS AND COMMUNICATIONS SECTION 23*rd February*, 1959.)

SUMMARY

An improved bandwidth has been obtained from a six-cavity pulsed S-band klystron amplifier by the use of a double-tuned output circuit. The output power was greater than 2·5 MW over a 6·3% bandwidth when the drive power was optimized at each wavelength, and greater than 2 MW over a 5·5% bandwidth with a constant drive of 25 watts. The power in the electron beam was 10 MW.

LIST OF SYMBOLS

d_1 = Inner door aperture.
d_2 = Outer door aperture.
I_1 = Driving current.
l = Length of second cavity.
P_e = Output power.
Q = Q-factor of cavity.
R_e = Output load resistance.
Z_{12} = Transfer impedance.

(1) INTRODUCTION

A recent paper by King[1] described a six-cavity klystron amplifier which gave a saturated output power greater than 2·5 MW over a band of 5·3%. The useful bandwidth was limited by the variation with frequency of both the saturated output power and the drive power required. This paper describes results obtained with a similar valve, modified to obtain a greater bandwidth.

The bandwidth set by the saturated output power is largely determined by the impedance/frequency characteristic of the output circuit. Beaver *et al.*[2] have shown that by using a double-tuned output circuit this bandwidth can theoretically be increased up to twice that of a single cavity. An output circuit of this type was designed and incorporated in the valve to be described.

The bandwidth of the drive/frequency characteristic depends on the resonant frequencies of the individual cavities, especially the intermediate or floating cavities. The useful bandwidth can be widened by increasing the stagger-tuning of the cavities, but at the expense of gain and regularity of the response. Further improvement might be obtained by increasing the number of cavities, but there is a limit to the extent to which this is useful,[3] and it has not been attempted, on the grounds of mechanical complexity. Increased stagger-tuning was made available by off-setting the tuning ranges of the intermediate cavities, following the pattern of frequencies used in the earlier valve.[1]

(2) DESIGN OF THE OUTPUT CIRCUIT

The original output circuit consisted of a rectangular cavity, coupled to the output waveguide through a pair of inductive doors, as in Fig. 1. The double-tuned circuit was formed by adding a further pair of doors to form a second cavity, as in Fig. 2. The length, l, of the second cavity was chosen to give the desired resonant frequency, final adjustment being made with

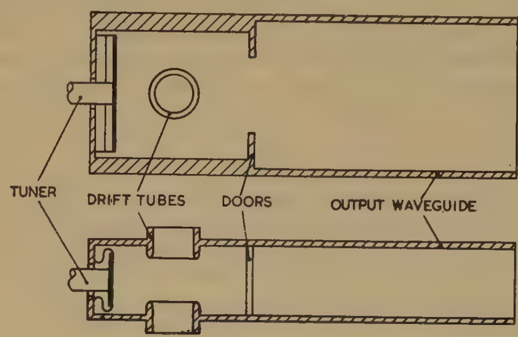


Fig. 1.—Single-cavity output system.

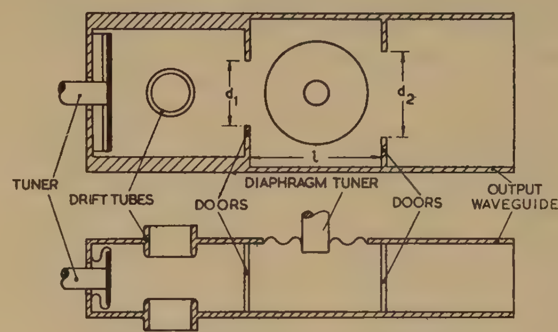


Fig. 2.—Double-cavity output system.

the diaphragm tuner. The capacitive stub was added to this to give a reasonable tuning range for a small movement of the diaphragm. The apertures, d_1 and d_2, between the doors, were chosen to give the appropriate frequency response, which was decided upon by the following method.

The power coupled into the external load from a bunched beam crossing the interaction gap is given by

$$P_e = \tfrac{1}{2}I_1^2|Z_{12}|^2/R_e$$

I_1 is the driving current at the gap, i.e. the effective r.f. current in the beam; R_e, the external load resistance; and Z_{12}, the transfer impedance, defined as the ratio of the voltage appearing across the load to the driving current. As the saturated driving current is largely independent of frequency, the saturated output power will vary with frequency as $|Z_{12}|^2$. In the case of a single-cavity system, Z_{12} is just the total impedance across the gap, and it is possible to calculate the value of the loaded Q-factor required to optimize the bandwidth to any desired power level.[3] Using a model of the output system, the coupling to the cavity can be adjusted to give this value. With a doubly-resonant system the form of Z_{12} is not so simple, and an experimental method of optimizing the bandwidth is desirable.

The frequency characteristic of Z_{12} for any output circuit can

Mr. Curnow and Mr. Mathias are at the Services Electronics Research Laboratory.

be investigated by feeding into the output waveguide a signal of constant amplitude. The voltage appearing across the gap will then be proportional to Z_{12} and the output voltage of a square-law crystal loosely coupled to the cavity will be proportional to P_e. A model of the double-cavity output system was made in which the door apertures were adjustable. The relationship between the signal obtained from the crystal and the output power to be expected from the klystron was established by removing the outer doors, so that the model represented the single-cavity system of which the performance was known. In this way, the signal level, corresponding to the value of Z_{12} needed to give 2·5 MW output, was determined. The two apertures were then adjusted to give the maximum bandwidth at that signal level. It was found that a bandwidth of just under 7% was obtainable.

(3) EXPERIMENTAL RESULTS

Apart from the output system and the increased stagger-tuning available, the construction of the klystron was identical to that of the earlier one.[1] It was operated at 120 kV, with a collector current of 82 amp and a beam interception of about 6%. The pulse length was 2 microsec and the pulse repetition frequency 100 pulses/sec. The cavity frequencies were adjusted to give a saturated output of more than 2·5 MW over as wide a band as possible. The saturated output and drive required are shown in Fig. 3. The maximum bandwidth to 2·5 MW was 6·3%, compared with 5·3% obtained with the single-cavity output valve.[1]

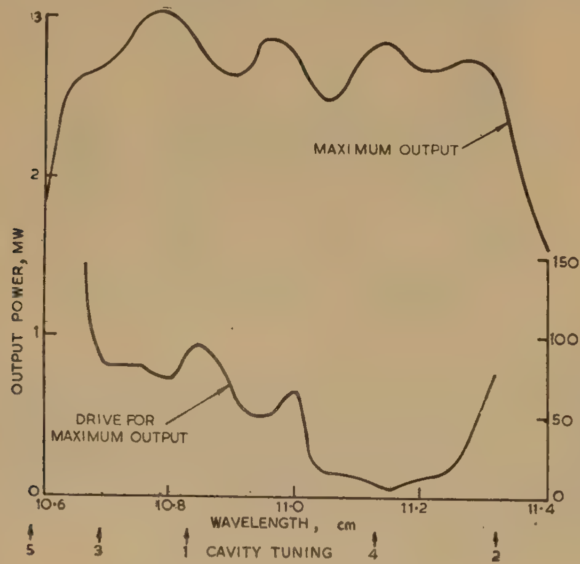


Fig. 3.—Saturated output characteristics.

Although this is a significant improvement, it is much less than that expected from the theory of Beaver *et al.*[2] The shape of the output characteristic of the earlier valve suggests that its output circuit was not truly singly-resonant (see Reference 1, Fig. 4). The best results were obtained by varying the loading of the output cavity with a capacitive matching stub in the output waveguide. This may have been acting as a second tuned circuit, giving some increase in bandwidth over the single-resonance value. Thus the failure of the double-tuned circuit to produce a greater increase in bandwidth may have been due to the enhanced performance of the earlier valve.

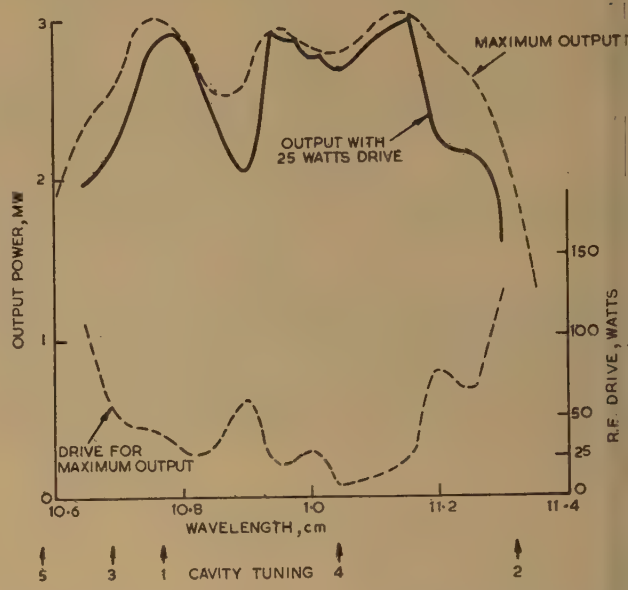


Fig. 4.—Output characteristics for a constant drive of 25 watts.

The cavity frequencies were also set to obtain the maximu 2 MW bandwidth with a constant drive. The results are sho in Fig. 4. More than 2 MW was obtained over a 5·5% ba with a constant drive of 25 watts.

Although the cavity tuning pattern was no longer limited the tuning ranges available, it was still not possible to set intermediate cavities to give a very flat drive over the increa bandwidth. It will be seen from Figs. 3 and 4 that the dr characteristic is tilted, compared with the response which mi be expected from the cavity resonances alone. The direction this tilt would indicate[1,3] that the cavities were too closely spa for optimum performance.

(4) CONCLUSIONS

By using a doubly resonant output circuit the bandwidth 2·5 MW has been increased from 5·3% to 6·3%. While i not suggested that the design tested was completely optimiz it seems unlikely that any great improvement could be m without employing even more complex circuits. The use bandwidth of the valve is limited by the drive required, improving this would involve adjusting the cavity separati and possibly increasing their number.

(5) ACKNOWLEDGMENT

This paper is published by permission of the Admiralty.

(6) REFERENCES

(1) King, P. G. R.: 'A 5% Bandwidth 2·5 MW S-Band Klystr *Proceedings I.E.E.*, Paper No. 2624 R, May, 1958 (10 Suppl. 12, p. 813).
(2) Beaver, W. L., Jepson, R. L., and Walter, R. L.: 'W band Klystron Amplifiers', *Institute of Radio Engin Wescon Convention Record*, 1957, **1**, Part 3, p. 111.
(3) Curnow, H. J.: 'Factors influencing the Design of M Cavity Klystrons', *Proceedings I.E.E.*, Paper No. 263 May, 1958 (**105** B, Suppl. 12, p. 855).

[The discussion on the above paper will be found on page 492.]

621.317.337

The Institution of Electrical Engineers
Paper No. 2847 E
Feb. 1959
©

A METHOD FOR THE MEASUREMENT OF VERY HIGH Q-FACTORS OF ELECTROMAGNETIC RESONATORS

By F. H. JAMES.

(*The paper was first received* 30*th September, and in revised form* 1*st December*, 1958. *It was published in February*, 1959, *and was read before the* ELECTRONICS AND COMMUNICATIONS SECTION 23*rd February*, 1959.)

SUMMARY

The paper outlines the problem of the measurement of the Q-factors of resonators for the C.E.R.N. linear accelerator. The Q-factors involved are of the order of 10^5, and the resonant frequencies are in the range 200–300 Mc/s. The methods considered are briefly discussed, and the one which seems most suitable employs sinusoidal amplitude modulation. The accuracy depends on a low-frequency phase measurement, and, for this reason, accurate knowledge of the power law of the detector is not required. The modulation component undergoes a phase shift owing to the high Q-factor of the resonator. When this attains a magnitude of $\pi/4$ rad, the Q-factor is given by one-half of the ratio of the resonant frequency to the modulation frequency. Two alternative procedures are described. A mathematical treatment of the theory and some experimental results are included in Appendices.

LIST OF SYMBOLS

v = Instantaneous voltage.
V_0 = Peak voltage.
i = Instantaneous current.
I_0 = Peak current.
R = Resistance.
C = Capacitance.
Z = Impedance.
t = Time.
ω = Angular frequency.
ω_0 = Angular frequency for resonance.
ω_m = Angular frequency of modulation.
$\omega_d = \omega - \omega_0$.
m = Degree of modulation.
ϕ = Phase difference of modulation component.
ψ = Phase difference of carrier component.
$\mathscr{I}$ = Imaginary part.
$\mathscr{R}$ = Real part.
$a_1 = 2Q\omega_d/\omega_0$.
$a_2 = 2Q\omega_m/\omega_0$.

$$\tan\gamma = \frac{a_2}{1 + a_1^2(1 + a_1^2 - a_2^2)/(1 + a_1^2 + a_2^2)}$$

$$\tan\gamma' = \frac{2a_2}{a_2^2 - a_1^2 - 1}$$

$$\tan\alpha = \omega_m RC$$

$$A = 1 + (a_1^2 - a_2^2)^2 + 2(a_1^2 + a_2^2)$$

(1) INTRODUCTION

At the laboratories of the European Organisation for Nuclear Research in Geneva, Switzerland, a 25 GeV proton synchrotron is being constructed. The injector for this high-energy particle accelerator will be a proton linear accelerator, which will accelerate protons from an initial energy of 500 keV to a final energy of 50 MeV. Three resonant cavities, operating in the E_{010}-mode of oscillation, will be used to effect this acceleration. The first of these cavities is 5·5 m in length and 1·08 m in diameter; without its drift tubes, which, in resonant linear accelerators, shield the accelerated protons from the reverse phase of the radio-frequency accelerating field, the resonant frequency is 212·8 Mc/s. The second and third cavities are 12 and 11·2 m in length, and 0·93 and 0·82 m in diameter, respectively. In the absence of the drift tubes, the resonant frequencies are 246·9 and 281·3 Mc/s. When the drift tubes are inserted in all three cavities, the common resonant frequency is 202·5 Mc/s. For each cavity a measurement of the Q-factor must be made under three separate experimental conditions, first without the drift tubes, secondly with the drift tubes, and finally with both the drift tubes and the high-power coupling loop. This final measurement is made in order to determine the loaded Q-factor, so that corrections can be made to the size of the loop if the desired coupling between the resonator and the coaxial feeder is not obtained. It is important that any measurement of the Q-factor be made in a short period of time, for the following reasons. Taking the resonant frequency as 200 Mc/s and the Q-factor as 10^5, the width of the resonance curve at the half-power points is only 2 kc/s. Since the temperature coefficient of the resonant frequency is approximately 4 kc/s per deg C, the measurement time must be short compared with the time for the mean temperature of the resonator to change by 0·01° C. Relatively high accuracy in these measurements is required, since a subsequent investigation of a low value of measured Q-factor may involve high costs and a loss of installation programme time. It would be even more embarrassing to make the discovery, after the drift tubes had been fitted, that the previous measurement had been optimistic, for it would then be necessary to remove all the drift tubes from the cavity and repeat the measurements.

(2) THE METHODS CONSIDERED

(2.1) Measurements Involving Slow Procedures

Methods* involving slow procedure are intended to include what could be considered as precision measurements, in order to satisfy the criterion of accuracy.

(*a*) A variable-frequency oscillator is used as a source of r.f. power, and is coupled to the resonator by some convenient system such as a loop. Observations are then made of the d.c. variations of a rectified r.f. signal from the resonator, which have been produced by measured changes in the oscillator frequency.

(*b*) A fixed-frequency oscillator is used, and the changes in frequency are effected by altering the resonant frequency of the resonator. Observations are made as in method (*a*), but a frequency calibration is required for the tuning mechanism which is incorporated in the resonator.

Mr. James is with the European Organisation for Nuclear Research, Geneva, Switzerland.

* An interesting method depending on r.f. phase measurement,[4] for which a high degree of accuracy is claimed, became known to the author subsequent to the consideration of the methods which are given above.

(*c*) To avoid the necessity for knowing the relationship between the r.f. power which is applied to the rectifier unit and the direct current which is caused to flow, the rectified current is maintained at a constant level. This is made possible by the inclusion of an r.f. attenuator. Observations are now made of the variations in the setting of the attenuator for measured changes in either the oscillator frequency or the resonant frequency of the resonator.

(*d*) A slotted line is used to make v.s.w.r. measurements for frequencies in the vicinity of the resonant frequency.[6]

(*e*) By measuring the magnetic field at the walls of the resonator, the theoretical Q-factor may be determined.[1]

(2.2) Measurements Involving Fast Procedures

Measurements involving fast procedures, with one exception, may not be as precise as those previously mentioned, but they would satisfy the criterion of rapidity of procedure.

(*a*) An f.m. oscillator is used as the source of power, and a resonance curve may be displayed on the screen of an oscillograph. An auxiliary r.f. oscillator may be used to provide a frequency calibrating system.[2]

(*b*) A pulsed a.m. oscillator is used, and, in this case, a measurement is made of the time required for the detected pulse from the resonator to decay to a measured fraction of the original amplitude.

(*c*) A sinusoidal a.m. oscillator is used, and a measurement is made of the phase difference between the modulation component of a signal which has passed through the cavity and a suitable reference signal.

(2.3) Comments on the Methods Considered

In our experience the methods of Section 2.1 can give only a limited accuracy. This may be expressed in terms of the probable errors encountered in typical results. For method 2.1(*a*), a probable error of $\pm 5\%$ could be achieved, but a more precise measurement required a very great effort. Since methods 2.1(*b*), (*c*) and (*d*) also involve a determination of experimental points from which a resonance curve can be constructed, it may be concluded that none of them will give a better result. Method 2.1(*e*) was considered because all the necessary apparatus was readily available for measurements not connected with the Q-factor. The method would have some applications whilst the drift tubes were not inserted, but not otherwise. It would therefore seem necessary to avoid a slow-procedure method. The difficulty presented by the fast-procedure methods is primarily due to the errors of measurement which are introduced by the oscillograph, since all three fast-procedure methods can be carried out using an oscillograph as the measuring instrument. There are, however, two fast-procedure methods which rely on amplitude measurements to determine the resonator power level, and one, namely method 2.2(*c*), which relies on a phase measurement for this determination. If there should be any doubt as to the relationship between the resonator power level and the indicated amplitude measurement, it would seem preferable to employ method 2.2(*c*) in conjunction with a highly sensitive phase-difference detector. This phase measurement can be carried out in our application at a frequency of 1–3 kc/s, which is not inconvenient. Therefore this is the method which has been chosen as being the most suitable.

(3) SOME DETAILS OF METHOD 2.2(c)

A resonant cavity is analogous to a series-tuned lumped circuit, and therefore to obtain a theory for method 2.2(*c*), let us consider an amplitude-modulated wave which is applied to such a circuit. Let a voltage v which is given by the expression

$$v = V_0 \sin \omega t(1 + m \cos \omega_m t) \quad . \quad . \quad .$$

be applied to an impedance Z, which is that of the series tun[ed] circuit, and is given by the approximation

$$Z = R[1 + 2jQ(\omega - \omega_0)/\omega_0] \quad . \quad . \quad .$$

The current which flows is given by

$$i = \frac{V_0}{R} \sin \omega_0 t\{1 + m \cos \phi \, [\cos (\omega_m t - \phi)]\} \, .$$

assuming that $\omega = \omega_0$, i.e. that the circuit is in exact resonan[ce]. The phase constant, ϕ, is given by

$$\phi = \arc\tan 2Q\omega_m/\omega_0 \quad . \quad . \quad . \quad . \quad .$$

It is therefore clear that a phase difference exists between t[he] modulation components of the current and the voltage in t[he] simple case of a series tuned circuit. When this phase differen[ce] attains the value $\pi/4$ rad, the Q-factor is given by

$$Q = \frac{\omega_0}{2\omega_m} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad ($$

A more detailed analysis is given in Appendix 7.1, where it [is] shown that, when the circuit is not exactly tuned, a phas[e] modulation term is introduced into the expression for the curren[t]. This term reduces to zero when the circuit is tuned. In additio[n] it is shown that there can be an error in the phase shift of th[e] modulation component. However, the error is shown to be [in] the same sense for either direction of mistuning. Consequent[ly] the tuning of the circuit could be completed by an observatio[n] of the limiting value of this phase shift. For the series tune[d] circuit previously considered, the voltage and the current can b[e] more easily determined than for a resonator. With a resonato[r] it is found that a rectified signal which has been derived from th[e] power applied to the resonator can provide a reference phase f[or] the modulation component. Another rectified signal is derive[d] from the resonator, and the phase of this signal is found [to] exhibit a variation similar to that calculated for the phase of th[e] modulation component of the current in the series tuned circu[it]. Therefore the application of eqn. (5) to the determination of th[e] Q-factor of a resonator involves the accurate measurement of th[e] phase difference between the two signals described. In practic[e] two alternative procedures are possible. In the first, a detect[or] for $\pi/4$ rad is used, and the modulation frequency is varied un[til] the phase difference attains this value. In the second procedu[re] a 0 rad phase detector is used. The reference signal is passe[d] through a phase-shifting network whose characteristics can b[e] varied to obtain zero phase difference relative to the rectifie[d] signal from the resonator. Once this condition is obtained, th[e] phase difference is independent of the modulation frequenci[es] from zero to the limit where differences between the two rectifi[er] units become significant. The second procedure may be superi[or] to the first, in that harmonics of the modulating frequency a[re] not so serious. Both branches can discriminate against ha[r]monics to the same degree. Appendix 7.2 gives an analysis [of] an *RC* circuit of the type envisaged.

(4) CONCLUSIONS

By employing the low-frequency phase-comparison techniqu[e] drifts in frequency of either the source of r.f. power or t[he] resonator are indicated by a phase error. This can be correct[ed] before the experimental observation is made from which t[he] Q-factor is determined. No assumptions are necessary wi[th] respect to the characteristics of a rectifier, nor do small variatio[ns]

in the power level of the oscillator give rise to errors in the measured value of the Q-factor. Although other methods give adequate results for small resonators which are measured under good experimental conditions, the method is worth consideration either when greater accuracy is required or when the conditions for measurement are unfavourable.

(5) ACKNOWLEDGMENTS

The author wishes to acknowledge the contributions made by Messrs. J. Sharp and P. Mulder, who designed a suitable $\pi/4$ phase-detecting meter, and Mr. B. Agoritsas who assisted with the design of a calibrator for this phase meter, and who constructed all the measuring apparatus which was required. Members of the C.E.R.N. linear-accelerator group, and, in particular Dr. H. G. Hereward, helped with discussions. The author is grateful to the Director General of C.E.R.N. for granting permission to publish the paper.

(6) REFERENCES

(1) KITCHEN, S. W., and SCHELBERG, A. D.: 'Resonant Cavity Field Measurements' (University of California Radiation Laboratory Report No. UCRL; 1947).

(2) ADLAM, J. H., JAMES, F. H., and MULLETT, L. B.: 'A Method of Measuring the Q-Factor of an Electromagnetic Mode in a Resonator Employing a Frequency Modulated Klystron' (Atomic Energy Research Establishment Report No. G/R767).

(3) LAPORT, E. A.: 'Characteristics of Amplitude Modulated Waves', *R.C.A. Review*, 1937, **1**, p. 26.

(4) BOYD, W.: 'Resonant Cavity Measurement Techniques at 400 Mc/s' (Atomic Energy Research Establishment Report No. GP/R2184).

(5) TERMAN, F. E.: 'Radio Engineers Handbook' (McGraw Hill, 1943).

(6) SLATER, J. C.: 'Microwave Electronics' (Van Nostrand, 1954).

(7) APPENDICES

(7.1) The Theory for the Phase Shift of the Modulation Component

An equation for a sinusoidal amplitude-modulated wave may be written as

$$v = \mathscr{I}\left\{V_0\varepsilon^{j\omega t} + V_0\frac{m}{2}\left[\varepsilon^{j(\omega+\omega_m)t} + \varepsilon^{j(\omega-\omega_m)t}\right]\right\} \qquad (6)$$

and if this voltage is applied to a series tuned circuit, whose impedance is given by the approximate relation

$$Z = R[1 + 2jQ(\omega - \omega_0)/\omega_0] \quad . \quad . \quad . \quad (7)$$

the current is given by

$$i = \mathscr{I}\left\{\frac{V_0\varepsilon^{j\omega t}}{R[1 + 2jQ(\omega - \omega_0)/\omega_0]} + \frac{V_0}{R}\frac{m}{2}\left[\frac{\varepsilon^{j(\omega+\omega_m)t}}{1 + 2jQ(\omega - \omega_0 + \omega_m)/\omega_0} + \frac{\varepsilon^{j(\omega-\omega_m)t}}{1 + 2jQ(\omega - \omega_0 - \omega_m)/\omega_0}\right]\right\} \qquad (8)$$

In order to make a simplification, and to bring out more clearly that ω may not, in general, coincide with ω_0, we put

$$\omega - \omega_0 = \omega_d \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

Furthermore let

$$a_1 = 2Q\omega_d/\omega_0 \quad . \quad . \quad . \quad . \quad . \quad . \quad (10)$$

$$a_2 = 2Q\omega_m/\omega_0 \quad . \quad . \quad . \quad . \quad . \quad . \quad (11)$$

and also

$$I_0 = V_0/R \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (12)$$

Then

$$i = \mathscr{I}\left(\frac{I_0\varepsilon^{j(\omega_0+\omega_d)t}}{1 + ja_1}\left\{1 + \frac{m}{2}\left[\frac{(1 + ja_1)\varepsilon^{j\omega_m t}}{1 + j(a_1 + a_2)} + \frac{(1 + ja_1)\varepsilon^{-j\omega_m t}}{1 + j(a_1 - a_2)}\right]\right\}\right) \quad (13)$$

From this expression, after rationalizing the denominators,

$$i = \mathscr{I}\left(\frac{(1 - ja_1)}{(1 + a_1^2)}I_0\varepsilon^{j(\omega_0+\omega_d)t}\left\{1 + \frac{m}{2A}(1 + a_1^2 + a_2^2)\right.\right.$$
$$[(1 + a_1^2)(\varepsilon^{j\omega_m t} + \varepsilon^{-j\omega_m t}) + (a_1a_2 - ja_2)(\varepsilon^{j\omega_m t} - \varepsilon^{-j\omega_m t})]$$
$$\left.\left. - \frac{m}{A}(a_1a_2)[(1 + a_1^2)(\varepsilon^{j\omega_m t} - \varepsilon^{-j\omega_m t}) + (a_1a_2 - ja_2)(\varepsilon^{j\omega_m t} + \varepsilon^{-j\omega_m t})]\right\}\right) \quad (14)$$

where

$$A = 1 + (a_1^2 - a_2^2)^2 + 2(a_1^2 + a_2^2) \quad . \quad . \quad . \quad (15)$$

Hence

$$i = I_0\cos\psi\{\sin[(\omega_0 + \omega_d)t - \psi]\}\left(1 + \frac{m}{A}(1 + a_1^2 + a_2^2)\left\{\left[1 + a_1^2\frac{(1 + a_1^2 - a_2^2)}{1 + a_1^2 + a_2^2}\right]^2 + a_2^2\right\}^{1/2}\cos(\omega_m t - \gamma)\right)$$
$$+ I_0\sin\psi\{\cos[(\omega_0 + \omega_d)t - \psi]\}\left(\frac{m}{A}(1 + a_1^2 + a_2^2)a_2\left\{\left[1 - \frac{2(1 + a_1^2)}{1 + a_1^2 + a_2^2}\right]^2 + \frac{4a_2^2}{(1 + a_1^2 + a_2^2)^2}\right\}^{1/2}\sin(\omega_m t + \gamma')\right) \quad (16)$$

where

$$a_1 = \tan\psi \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (17)$$

$$a_2 = \tan\phi \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (18)$$

$$\tan\gamma = \frac{a_2}{1 + a_1^2(1 + a_1^2 - a_2^2)/(1 + a_1^2 + a_2^2)} \quad . \quad (19)$$

$$\tan\gamma' = \frac{2a_2}{a_2^2 - a_1^2 - 1} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (20)$$

Putting in the conditions for exact resonance,

i.e. $$\omega_d = 0 = a_1 = \psi \text{ and } \gamma = \phi \quad . \quad . \quad . \quad (21)$$

Then $$i = I_0\sin\omega_0 t[1 + m\cos\phi\cos(\omega_m t - \phi)] \quad . \quad (22)$$

But $$v = V_0\sin\omega_0 t(1 + m\cos\omega_m t) \quad . \quad . \quad . \quad . \quad (23)$$

which shows that the modulation components of the applied voltage and current have a phase difference of ϕ, and when $\phi = \pi/4$

$$Q = \omega_0/2\omega_m \quad . \quad . \quad . \quad . \quad . \quad . \quad (24)$$

When the tuning is not exact, there is an error in phase given by $(\gamma - \phi)$, and a phase-modulation term. However, the phase error depends on the square of the frequency error, so that the limiting value of the phase difference as determined by the phase meter should determine the exact tuning position.

(7.2) Phase Shift Caused by a Resistance and a Capacitance in a Series Circuit

Let a sinusoidal voltage $v = V_0 \cos \omega_m t$ be applied to a series RC circuit, and consider the voltage which is developed across the capacitance.

The voltage is given by

$$\mathscr{R}\left(\frac{1}{1 + j\omega_m RC} V_0 \varepsilon^{j\omega_m t}\right) \quad . \quad . \quad . \quad (25)$$

which, after rationalization of the denominator, and taking the real part becomes

$$\frac{(\cos \omega_m t + \omega_m RC \sin \omega_m t)V_0}{1 + \omega_m^2 R^2 C^2} \quad . \quad . \quad . \quad (26)$$

Let

$$\omega_m RC = \tan \alpha \quad . \quad . \quad . \quad . \quad . \quad . \quad (27)$$

The voltage which is developed across the capacitance is given by

$$\cos \alpha \cos (\omega_m t - \alpha) V_0 \quad . \quad . \quad . \quad . \quad (28)$$

If this expression is compared with the modulation component of the current which was given in eqn. (22), namely

$$m \cos \phi \cos (\omega_m t - \phi) I_0$$

it is then clear that, for $\phi = \alpha$, the phase and amplitude depe dencies are the same for both expressions.

Furthermore, if

$$\cot \alpha = 1 \quad . \quad . \quad . \quad . \quad . \quad . \quad (2$$

then

$$\omega_m = 1/RC = \omega_0/2Q \quad . \quad . \quad . \quad . \quad . \quad (3$$

(7.3) Some Experimental Results

For the first cavity, with the drift tubes, 14 measurements ga the result $Q = 62\,500$ with a probable error of 500. Anoth series of ten measurements for the first cavity, with the dr tubes, gave the result $Q = 62\,300$, with a probable error of 45

For the third cavity, without the drift tubes, 21 measureme gave the result $Q = 78\,000$ with a probable error of 1 400.

The second cavity was measured without the drift tubes befo the final form of the measuring equipment had been realize As a consequence the result is rather poor, and the result w $Q = 79\,000$ with a probable error of 3 000.

Nevertheless, this probable error was much smaller than th obtained using other methods for the Q-factor measurement.

DISCUSSION ON THE ABOVE FOUR PAPERS BEFORE THE ELECTRONICS AND COMMUNICATION SECTION, 23RD FEBRUARY, 1959

Mr. J. Dain: In Paper No. 2789 the author calculates the distortion of an f.m. signal arising in a klystron connected to a mismatched load. This effect should be considered in conjunction with the incidental amplitude modulation which arises when the klystron frequency is modulated. Thus, in principle, it is always possible to reduce the load pulling of the oscillator by incorporating a stabilizing cavity in the output circuit. Inevitably this increases the amplitude modulation so that the degree of stabilizing which is introduced must strike a compromise between the two types of distortion. Has the relative importance of the f.m. and the a.m. distortion been considered?

In Paper No. 2788 the authors describe a helix structure which they propose to use in a pulsed high-power travelling-wave tube. The helix has two disadvantages when it is constructed to propagate a wave with a phase velocity approaching the velocity of light, because the conductor tends to a straight wire running along the axis of the system. As a result considerable power flows external to the helix and the electric field tends to be purely transverse; both these effects reduce the coupling impedance presented to the beam by the structure. The structure shown in the paper avoids the high flow of power external to the helix by the addition of the conducting shield, but it does not alleviate the tendency of the electric field to become transverse. The coupling impedance would still be zero if the design was for a wave with a velocity equal to that of light. Other structures, such as the disc-loaded waveguide, do not have this restriction, although of course their bandwidth is more limited. Perhaps the authors would make some comment.

On the atomic scale Nature works in discrete, quantized steps; on the macroscopic scale she prefers to work gradually and avoid sudden changes as much as possible. I suggest this should be borne in mind when designing the output system of a high-power klystron. The r.f. energy carried by the beam should not be extracted at one resonator gap spanned by a voltage which is about equal to the direct beam voltage. Instead, for example, it should be removed at two resonator gaps spanned by voltages of suitable phase and each of magnitude approximately one-half the beam voltage. Theoretically this system should have a bandwidth twice that of the single cavity. The logical conclusion of this process is to increase the number of cavities even further and so, perhaps, arrive at an output system which is a compromi between a single-cavity circuit and a travelling-wave circuit.

I should like to draw Mr. James's attention to a paper Keith-Walker* dealing with the measurement of Q-factors excess of 10^6. The apparatus described included a neat frequen discriminator used in conjunction with an f.m. klystron. Th arrangement avoids stabilizing the oscillator frequency and resu in greater simplicity of the apparatus.

Mr. R. B. R.-Shersby-Harvie: The simple and elegant meth of measuring Q-factor described in Paper No. 2847 can regarded in another way. It can be seen that the resonator h a linear rate of change of phase in the response over the sma bandwidth. This means that a time delay can be expected in t outgoing signal from the ingoing, but, if the bandwidth of t signal is not too great, not much shape distortion will be expecte In fact, the shape of the modulation is not very importar Would it be possible in some applications, where it is not ve convenient to modulate sinusoidally, to use a pulse? F instance, if the response is plotted against time with a sha square pulse the usual response building-up and decaying obtained, which is a perfectly legitimate way of measuring Q decay time. On the other hand, if a more rounded pulse with narrower bandwidth were used, possibly a signal of the sar shape but delayed in time would be obtained. Has the auth carried out any experiments with pulses, or has he any ideas what would happen?

Dr. G. D. Sims: In Section 1 of Paper No. 2788 is is stated th the helix structure can be used in a travelling-wave tube of t order of 30% efficiency. Why are people still content to ma tubes with efficiencies as low as this? Fairly clearly, with lo level tubes these efficiencies are sufficient, but for handling hi power it is time the problem was examined further.

If the tube is 30% efficient, 70% of the d.c. input power wasted, of which some will be dissipated in the structure its but most will be left on the beam at the end of its transit throu the slow-wave structure. I wonder how much of this lat power is wasted, because by the time the end of the slow-wa

* KEITH-WALKER, D. G.: 'An Equipment for Measuring the Attenuation of L Loss Waveguide Transmission Lines', *Proceedings I.E.E.*, Paper No. 2838 R, Janu 1959 (106 B, Suppl. 13, p. 71).

structure is reached the beam has given up perhaps 50% of its total energy and is very much out of step with the travelling wave.

This idea is familiar in linear accelerator work. In an accelerator the phase velocity is always arranged so that the particles in the beam stay in step with the wave, right down the length of the tube. With a 200 kV beam, if the beam gives up 50% of its power in travelling through the tube, a 12% reduction in velocity will take place between the input and output ends. Could not graded phase-velocity structures be used for travelling-wave tubes similar to those used in linear accelerators?

The authors wanted to make use of as many parts of the original klystron as possible in making up their travelling-wave tube. Why did they not retain the buncher resonator from the klystron and send a ready-bunched beam into the helix? Might this not also have resulted in some increase in efficiency? Admittedly the bandwidth of the valve would be reduced, but there is no reason why a useful tube of intermediate bandwidth should not be made, provided that a low-Q buncher is used.

Mr. N. D. West: At the Atomic Energy Research Establishment, we are building a 15 MeV linear accelerator as an injector for a 7 GeV proton synchrotron. The accelerator will consist of a single resonant cavity, similar in structure to those described in Paper No. 2847, and will resonate at 115 Mc/s with a Q-factor of about 10^5. We also wish to make accurate Q-factor measurements, and have decided to adopt the author's method as suggested by him earlier in a C.E.R.N. internal report.

We have constructed suitable apparatus, employing the RC phase-shift network, in which zero phase difference between the two a.f. signals is detected using a long-tailed-pair difference amplifier. The apparatus has been tried out on a test cavity having a Q-factor of about 40 000. Varying both the modulation frequency and the ratio $R : C$ in the phase-shift network, several readings of Q-factor were obtained, all of which were within 1% of the mean value.

It was found possible to take readings very quickly, so that drift in the frequency of the r.f. oscillator or in the cavity resonant frequency created no difficulty. As little circuitry as possible was employed in order to avoid unwanted phase shifts, and by using the phase-shift network method no measurement of audio frequency is required, since the Q-factor is given simply in terms of calibrated values of R and C.

We have found this method of Q measurement very successful at this frequency and Q value, and I would like to ask Mr. James if he feels that this method has any application at other frequencies, in particular, in the microwave region.

Mr. W. E. Willshaw: It is clear from the presentation of Paper No. 2789 that the author deliberately simplifies the problem as far as he can to make the results accessible to analysis. In Section 2 he ignores any difference there may be between the static and dynamic Riecke-diagram characteristics of the klystron. Not to do this involves considerable difficulty, but in applying this analysis to a practical case, has the author any views on how significant the difference might be? One is concerned here with a self-oscillator valve with a finite build-up time, and the change of frequency due to the reflections of the load system cannot take place instantaneously. How significant is this effect?

Has the author any experimental data relating the analysis he has done to modulation characteristics in actual reflex klystrons?

Mr. D. T. Swift-Hook: In Paper No. 2788, the beam power of 20 MW associated with the bandwidth of a helix offers the possibility of a considerable advance in the travelling-wave-tube field. What applications do the authors envisage for such a bandwidth at this power level?

While it is true that helix-type travelling-wave tubes known at the present time are relatively low-power devices when compared with the 20 MW aimed at in the paper, they are not quite as low-power as the one cited by the authors which has only a 30-watt beam. Tubes are commercially available in this country at S-band with continuous beam power of more than six times the power quoted. At the recent convention on microwave valves a pulsed X-band helix tube was described having a beam power of 10 kW and an r.f. power of more than 1 kW. An experimental valve has been made with a beam power of 60 kW; this valve was for a c.w. v.h.f. television transmitted centred on 750 Mc/s and had a power output of 14 kW. While still not in the megawatt class, I think these figures give a better idea of the state of existing valves.

Concerning the large wire diameter (0·16 of the helix diameter), sheath helix analyses have already been shown to be useful for wire diameters as large as 0·26 of the helix diameter, as in the X-band pulsed tube already mentioned.

Can the authors say what limitations they expect due to backward-wave oscillations on their unstrapped helix, and how far the bandwidth will be reduced by the strapping required to reduce these spurious oscillations?

Can the authors give the latest information on the hot valve? Has a beam yet been put down this very interesting structure?

Mr. B. B. Jacobsen (*communicated*): The static approach used in Paper No. 2789 is approximate in that it appears not to take into account the fact that at least the outer sideband components of the wave reflected from the load will be reflected back from the oscillator, which one would imagine to have a very high reflection factor at frequencies removed from the instantaneous carrier frequency.

Figs. 4–7 would have been much more meaningful if they had been expressed in terms of the logarithm of the reflection factor rather than in terms of the v.s.w.r. of the load.

THE AUTHORS' REPLIES TO THE ABOVE DISCUSSION

Mr. D. T. Gjessing (*in reply*): In reply to Mr. Dain, the frequency pulling described in the paper will no doubt give an additional amplitude modulation. However, for all practical applications of a klystron as a frequency modulator, an operating point in the Riecke diagram will be so chosen that the change in output power resulting from the pulling effect is small—certainly less than 10%. The amplitude limiter on the receiver will then remove this amplitude modulation before the f.m. signal is detected. If, however, the f.m. klystron is feeding an a.m. distortion signal into, for example, a travelling-wave tube, this may lead to phase distortion. This distortion source is, however, easily neutralized by connecting the klystron directly to the power amplifier, thus nullifying both types of pulling distortion. I do not think it is commendable to use a stabilizing cavity as a means of reducing the pulling effect of a f.m. klystron as this will have a highly undesirable effect on the modulation properties of the klystron if the coupling of the cavity, its Q-factor and resonance frequency are not very carefully chosen.

In reply to Mr. Willshaw, I find it difficult to visualize how the static Riecke diagram can be different from the dynamic one for cases of practical significance. It is well known that, as far as the dynamic modulation characteristics are concerned, the amplitude modulation begins to depart from static behaviour when times comparable with the decay time of the loaded cavity or frequencies comparable with the bandwidth of the cavity are involved. As regards frequency modulation with small deviations about the maximum power point, there is no time rate of change of energy stored in the circuit and one would not expect

any difference between the static and the dynamic case. Even though this argument is related to the control of the electron beam by a voltage applied to an electrode, I see no reason why the same argument could not be applied to the case where the electron beam is affected by a field resulting from a wave reflected back from a mismatched load. I have performed no comprehensive experimental investigation with the view to confirm the results derived in the paper.

Contrary to Mr. Jacobsen's opinion, I do not believe that the outer sideband components of the reflected wave impingent upon the oscillator will be reflected back to an extent appreciably greater than the centre frequency component. It is easy to show that, whether the klystron cavity is greatly undercoupled or overcoupled, the ratio of the half-power v.s.w.r. to the v.s.w.r. at resonance is 2 for the loaded cavity. This means that for a typical Q-factor of 100 at 4 000 Mc/s the v.s.w.r. for sidebands as far out as 20 Mc/s is not more than twice that at the carrier frequency.

Messrs. G. W. Buckley and **J. Gunson** (*in reply*): Fig. 4 supports the power-flow comments of Mr. Dain and shows the division of the power flow both inside and outside the helix. It is seen, however, that the internal power flow is by no means negligible.

The fundamental transverse components of the electric field contain the factor β/γ, whereas the axial component does not. Whilst β/γ tends to infinity as the phase velocity approaches the velocity of light, when $v_1 = 0\cdot67c$, β/γ is only about $1\cdot3$. Hence the resultant electric field is far from being purely transverse.

It is true that the coupling impedance of the helix is algebraically zero when the supported wave has a phase velocity equal to that of light. However, this is scarcely a real point against the helix, as it would have degenerated into a straight conductor parallel to the electron beam. (It is worth noting that under this condition, the energy of the electrons is algebraically infinite.) There must be some practical limit to the magnitude of the phase velocity of helix which is useful in a travelling-wave tube. The higher the phase velocity, the higher must be the voltage of the electron gun. This implies physically larger electron guns than are known at present if a perveance of 10^{-6} is to be preserved. Electron guns have worked at 400 kV, but they have also been the seat of unwanted oscillations due to feedback through the anode drift tube. Accordingly, it would seem that for a practical tube the limiting factor is not necessarily the helix.

We share Dr. Sim's sentiments about low conversion efficiencies. These apply equally to the megawatt types of klystron which have efficiencies of the order 30–40%. As yet, however, we are unaware of a practical method of increasing this efficiency although several schemes have been put forward.

The paper records the results of an investigation of the 'cold' properties of a particular helix structure. The next step is to examine the properties of the structure in conjunction with an electron beam. If travelling-wave amplification resulted with an efficiency of about 30%, we would be delighted. Only at this stage in the programme would we be justified in considering modifications to the uniform helix. Grading the phase velocity to suit the slower electron beam seems to be a promising idea and was suggested by the late Dr. A. W. Aikin a few years ago.

The behaviour of the structure in conjunction with an electron beam is the only safe way of determining whether backward wave oscillations exist. We are hopeful of success with the strapped structure as the currents in adjacent helix members are tied together. In addition, we believe the fundamental mode to be basically unchanged and think the higher modes have been considerably modified.

We thank Mr. Swift-Hook for the present-day information on travelling-wave tubes and are pleased to substitute a 180-wa[tt] beam for one of 30 watts. However, the text of the paper wa[s] prepared before the 1958 Microwave Convention and we we[re] unaware at that time of Mr. Swift-Hook's X-band helix tub[e]. We are encouraged by its success and are glad to know th[e] magnitude of the ratio of wire to helix diameter. Naturall[y] we are curious to know the magnitudes of the correspondi[ng] radial propagation constant and the helix diameter.

Messrs. H. J. Curnow and **L. E. S. Mathias** (*in reply*): Th[e] possibility of using two output cavities, coupled to each oth[er] and to the beam, was considered before adopting the schem[e] described. Such an arrangement would at first sight give th[e] same bandwidth, with possibly higher efficiency, but at th[e] expense of greater mechanical complexity. Moreover, th[e] coupling of the cavities by the electron beam complicates th[e] situation, and may even lead to oscillation.

In general, cross-breeds between klystrons and travelling-wav[e] tubes are unlikely to be successful, as the high-signal-lev[el] bunches produced by one system are not suitable for injecti[on] into the other.

Mr. F. H. James (*in reply*): I am very grateful to Mr. J. Dai[n] for bringing Mr. Keith-Walker's paper to my notice. There a[re] two criteria which may be used in order to determine th[e] applicability of the method which is described in my pape[r]. The first is the ratio of the resonant frequency to the Q-facto[r]. This ratio determines the bandwidth over which the inpu[t] channels of the phase-difference detector are identical, withi[n] the limits of the desired accuracy. For example, consider [a] resonant frequency of 25 Gc/s and a Q-factor of 10^7; this band[-] width requirement is the same as that for a resonant frequenc[y] of 200 Mc/s and a Q-factor of 8×10^4. The second criterio[n] is the facility with which one can amplitude modulate a[n] oscillator. Clearly at 200 Mc/s the second criterion is easil[y] satisfied. The types of oscillator which are used in the micro[-] wave region are more difficult to amplitude modulate withou[t] attendant frequency modulation. The analysis which was give[n] in the paper may easily be extended to include the effects o[f] parasitic frequency and phase modulation, and from this analys[is] it is possible to determine the permissible parasitic modulatio[n] indices when the $\pi/4$ rad detector is used. This difficulty is [a] possible reason for preferring other methods at microwa[ve] frequencies.

Using the $\pi/4$ rad detector the bandwidth is restricted only b[y] the cavity signal branch of the system, so that, if harmonics o[f] the modulating frequency are present, distortion of the origin[al] modulation signal will occur only in this branch. Consequent[ly] there will be an error in the measured Q-factor. If, however, th[e] zero rad phase detector is used and an *RC* circuit is included in th[e] reference signal branch, the bandwidths and associated dis[-] tortions of the two branches are equalized by the correct adjus[t-] ment of the *RC* network. Thus the second alternative which [is] outlined in the paper, and has been used by Mr. West, is certain[ly] the better procedure, since the phase detector is simpler and th[e] harmonic content of the modulation is not important.

In reply to Mr. R.-Shersby-Harvie, our experiments have bee[n] confined to the use of approximately sinusoidal amplitu[de] modulation. His example of a sharp pulse or a moderate[ly] rounded one may be taken as a modulating signal with hig[h] harmonic content, for which the zero rad phase-difference detect[or] can be used in conjunction with an *RC* circuit. In this case t[he] time delays and attenuations of all harmonics due to the Q-fact[or] of the cavity are simultaneously equalized by the *RC* circuit [in] the reference branch; the final shapes of the output pulses a[re] the same, and the pulse repetition frequency should not aff[ect] the phase equality between the two branches, provided that t[he] *RC* circuit is in correct adjustment.

621.372.822:621.317.373

The Institution of Electrical Engineers
Paper No. 2854 E
Sept. 1959
©

PHASE MEASUREMENTS THROUGH TAPERED JUNCTIONS

By L. LEWIN, Associate Member.

(*The paper was first received* 4*th November, and in revised form* 20*th December*, 1958.)

SUMMARY

If a tapered section is suitably matched at the transitions between two uniform waveguides, the electrical length of the taper is usually taken to be its physical length. It is shown here that a small correction is necessary on account of the cylindrical nature of the wave which propagates in the taper region, and a simple expression for the correction is found for a rectangular waveguide in the dominant mode.

(1) INTRODUCTION

The measurements of phase or position of standing-wave patterns in a waveguide are usually conducted in the guide itself, through the use of a slotted line. Occasionally, however, it may be convenient to perform the measurements in a waveguide of different dimensions, because of a design involving non-standard guide, the availability of an instrument in another guide size, or for some other reason. Under these circumstances it is normal to use a gradual taper from the one waveguide to the other, the phase measurements then involving the length of tapered transition. Owing to reflections at the junction of the tapered section with the uniform waveguides, matching elements must be used if errors are to be avoided. Both in checking such matching devices and in the measurements themselves, the effective electrical length of the taper is needed, and this should, of course, agree with the deductions from measurements using a short-circuiting plunger in the waveguide under test. Any discrepancy requires explanation before the arrangement can be considered reliable, and for this purpose the effective length of the taper must be known.

It is usually assumed that the electrical and physical lengths of the taper are the same, the taper being designed to be of such a length that the discrepancy due to the difference between the slant and axial lengths can be ignored, and at the same time providing a fairly good match at the transition. However, the propagation of an electromagnetic wave in a taper involves cylindrical functions whereby the electric and magnetic components turn out to be not exactly in phase, especially near the origin. This gives rise, on the one hand, to a slight mismatch at the transition from uniform to tapered waveguide,* and on the other to a small alteration in the effective length of the taper. The purpose of the paper is to explain how this comes about, and to give a formula for the correction. To this end we will analyse the simplest possible arrangement, namely two parallel plates tapering (linearly) from a uniform section. The tapered section will continue either to infinity (the matched case) or will be short-circuited by a cylindrical end-cap (the short-circuited case). A cylindrical rather than a flat end-cap is used to avoid introducing complications due to phasing errors arising from the difference between the slant and axial lengths of the taper: as already explained, these will, in practice, be small enough to ignore.

(2) THE ELECTROMAGNETIC FIELD

Fig. 1 shows two parallel plates, of spacing b, which taper from $x = 0$ with a half angle ϕ_0. The effective vertex of the taper is at $x = -r_0$, from which point the radial vector r is measured. At $r = r_1$, corresponding to $x = l$, the taper length, the taper is terminated with a cylindrical end-cap. The field† in the uniform section may be written

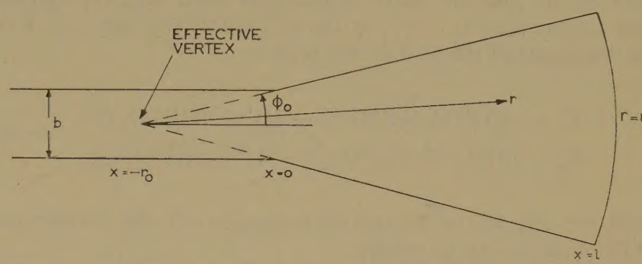


Fig. 1.—Uniform taper from parallel-plate region, terminated by cylindrical end-cap.

$$\left.\begin{aligned} E_z &= \varepsilon^{-jkx} + R\varepsilon^{jkx} \\ H_y &= \varepsilon^{-jkx} - R\varepsilon^{jkx} \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad (1)$$

where $k = 2\pi/\lambda$ and R is the voltage reflection coefficient. (In a waveguide of finite height it is sufficient to replace k by $k' = 2\pi/\lambda_g$ at the end of the analysis.)

Consider first the case of a matched termination. Then there will be no reflections from the right of the junction, the field in that region being given by

$$\left.\begin{aligned} E_\phi &= T\mathrm{H}_1^{(2)}(kr) \\ H_y &= jT\mathrm{H}_0^{(2)}(kr) \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad (2)$$

Here T is the transmission coefficient and $\mathrm{H}_{0,1}^{(2)}$ are the Hankel functions of the second kind of zero and first order, representing outgoing waves. They possess asymptotic expansions which, for large values of argument (small taper angles), can be written approximately as

$$\left.\begin{aligned} \mathrm{H}_0^{(2)}(z) &\simeq (2/\pi z)^{1/2} \exp\left[j\pi/4 - j(z - 1/8z)\right] \\ \mathrm{H}_1^{(2)}(z) &\simeq (2/\pi z)^{1/2} \exp\left[j3\pi/4 - j(z + 3/8z)\right] \end{aligned}\right\} \quad (3)$$

The functions of the first kind, with superscript unity and representing incoming waves, differ from eqn. (3) only in having the sign of j reversed.

On equating the fields from eqns. (1) and (2) at the centre of the transition* the following equations are derived:

$$\left.\begin{aligned} 1 + R &= T\mathrm{H}_1^{(2)}(kr_0) \\ 1 - R &= jT\mathrm{H}_0^{(2)}(kr_0) \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad (4)$$

From these the admittance at the junction is deduced as

$$\begin{aligned} Y = \frac{1-R}{1+R} &= \frac{j\mathrm{H}_0^{(2)}(kr_0)}{H_1^{(2)}(kr_0)} \\ &\simeq \varepsilon^{j/2kr_0} \quad \text{from eqn. (3)} \\ &\simeq 1 + j/2kr_0 \quad . \quad . \quad . \quad . \quad . \quad (5) \end{aligned}$$

* LEWIN, L.: 'Advanced Theory of Waveguides' (Iliffe, 1951), p. 115.

† Since only relative values are of concern in the paper, it is convenient to take units such that the wave impedance of free space is unity.

Written contributions on papers published without being read at meetings are invited for consideration with a view to publication.
Mr. Lewin is with Standard Telecommunication Laboratories, Ltd.

Hence the effect of the junction on the admittance is to insert a parallel susceptance $j/2kr_0$. This can be tuned out by means of a parallel inductance of normalized value $2kr_0$, and it is assumed that this is done. Failure to do so will involve phase errors due to the mismatch, while matching by means of a series inductance, which is also possible, introduces a T-network at the junction, also with a phase change.

Reverting now to the short-circuited case, the field to the right of the junction must be modified from eqn. (2) to give zero electric field at $r = r_1$ at the cylindrical end-cap. If A is the amplitude of the field in this region,

$$\left.\begin{aligned} E_\phi &= A[\mathrm{H}_1^{(2)}(kr)\mathrm{H}_1^{(1)}(kr_1) - \mathrm{H}_1^{(1)}(kr)\mathrm{H}_1^{(2)}(kr_1)] \\ H_y &= jA[\mathrm{H}_0^{(2)}(kr)\mathrm{H}_1^{(1)}(kr_1) - \mathrm{H}_0^{(1)}(kr)\mathrm{H}_1^{(2)}(kr_1)] \end{aligned}\right\} \qquad (6)$$

With the help of the asymptotic expansions (3) the admittance at the junction can be written

$$Y \simeq \frac{-j\cos\left[\left(kr_1 + \frac{3}{8kr_1}\right) - \left(kr_0 - \frac{1}{8kr_0}\right)\right]}{\sin\left[\left(kr_1 + \frac{3}{8kr_1}\right) - \left(kr_0 + \frac{3}{8kr_0}\right)\right]} \quad . \quad (7)$$

Now, if the argument of the cosine is written in the form

$$\left(kr_1 + \frac{3}{8kr_1}\right) - \left(kr_0 + \frac{3}{8kr_0}\right) + \frac{1}{2kr_0}$$

and the cosine is expanded, in relation to the small quantity $1/2kr_0$, by means of the formula $\cos(x + \delta) \simeq \cos x - \delta \sin x$, then eqn. (7) for the admittance can be written

$$Y \simeq -j\cot(kl') + j/2kr_0 \quad . \quad . \quad . \quad . \quad (8)$$

where

$$\begin{aligned} kl' &= \left(kr_1 + \frac{3}{8kr_1}\right) - \left(kr_0 + \frac{3}{8kr_0}\right) \\ &= kl - \frac{3(1/r_0 - 1/r_1)}{8k} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \end{aligned}$$

since $r_1 - r_0 = l$, the physical taper length.

On comparison of eqns. (8) and (5) it is seen that the equivale circuit is a reactance, representing the reactive effect of t junction, in parallel with a short-circuited line of phase ang given by eqn. (9).

For a fixed frequency the tapered line accordingly appea shorter than the physical length by an amount

$$\delta l = \frac{3(1/r_0 - 1/r_1)}{8k^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad (1$$

However, if it should be the difference in phase which is interest as the frequency varies, differentiation of eqn. (9) wi respect to k (or $1/k$) is required, leading to the introduction of minus sign, so that the line appears longer by the same amou

(3) EXAMPLE

As an example to indicate the orders of magnitude which ma be involved, consider a $3\frac{1}{2}$ in taper from waveguide $\frac{3}{4}$ in wide one $\frac{1}{4}$ in wide at an operating wavelength giving $\lambda_g = 4$ in. F this arrangement the difference between axial and slant lengt is $\frac{1}{2}(1/3 - 1/8)^2/3\frac{1}{2} = 0{\cdot}006$ in, and is probably negligible. (Th would be a reasonable taper size to use in practice, althou the junctions would require matching with diaphragms.)

From the geometry of Fig. 1 we get, in the present cas $r_0 = 2{\cdot}1$ in and $r_1 = 5{\cdot}6$ in.

Hence, from eqn. (10) we find $\delta l = 0{\cdot}045$ in. This is not large error, but it is sufficiently important to be taken in account when performing accurate measurements. Failure do so would cause a discrepancy in relating short-circuit measur ments to expectations. The occurrence of a field-pattern mir mum apparently a short distance behind a physical short-circu was, in fact, the starting-point which led to the prese investigation.

PROCEEDINGS OF THE INSTITUTION OF ELECTRICAL ENGINEERS

Part B. ELECTRONIC AND COMMUNICATION ENGINEERING (INCLUDING RADIO ENGINEERING), SEPTEMBER 1959

CONTENTS

Declaration on Fair Copying.—Within the terms of the Royal Society's Declaration on Fair Copying, to which The Institution subscribes, material may be copied from issues of the *Proceedings* (prior to 1949, the *Journal*) *which are out of print and from which reprints are not available.* The terms of the Declaration and particulars of a Photoprint Service afforded by the Science Museum Library, London, are published in the *Journal* from time to time.

Bibliographical References.—It is requested that bibliographical reference to an Institution paper should always include the serial number of the paper and the month and year of publication, which will be found at the top right-hand corner of the first page of the paper. This information should precede the reference to the Volume and Part.

Example.—SMITH, J.: 'Reflections from the Ionosphere', *Proceedings I.E.E.*, Paper No. 4001 R, December, 1954 (102 B, p. 1234).

Published by The Institution, Savoy Place, London, W.C.2. *Telephone:* COVent Garden 1871. *Telegrams:* 'Voltampere, Phone, London.'
Printed by Unwin Brothers Limited, Woking and London.